

# AJS

American  
Journal  
of Sociology  
Volume 100 Number 4  
January 1995

*Capital and Social Structure in Cultural Fields*—Anheier,  
Gerhards, and Romo

*A Feminist Critique of Durkheim's Deviance and  
Suicide*—Lehmann

*Models for Describing the Structure of Sex  
Segregation*—Charles and Grusky

*Vote Turnout of 19th Amendment Women*—Firebaugh and  
Chen

*Class Mobility and Political Preferences*—De Graaf,  
Nieuwbeerta, and Heath

*Social Capital and Controlling Right-Wing Extremism in Berlin  
Youth*—Hagan, Merkens, and Boehnke

The University of Chicago Press

MARTA TIENDA, Editor

JIM DAVIS, FRED KNISS, SUSAN E. MAYER, JEFFREY MORENOFF, EMILIO PARRADO, and ROBERT J. SAMPSON, Associate Editors

MARY C. BRINTON, Associate Editor, Book Review

GRACE KAO, Assistant Book Review Editor

CHAN-UNG PARK, STEPHEN B. PLANK, MICHAEL J. REAY, DAVID A. SMILDE, and ALFORD A. YOUNG, JR., Book Review Editorial Board

SUSAN ALLAN, Managing Editor

LAUREN FORD-TAGGART and CLAIR JAMES, Editorial Assistants

**Consulting Editors** PAUL D. ALLISON • EDWIN AMENTA • PETER BEARMAN • ROBERT J. BURSİK, JR. • JOHN CASTERLINE • NANCY DAVIS • NANCY DENTON • CLAUDE FISCHER • ROBERTO FRANZOSI • ADAM GAMORAN • JOHN R. HALL • MICHAEL T. HANNAN • MICHAEL HECHTER • DOUGLAS HECKATHORN • BARBARA HEYNS • CHARLES HIRSCHMAN • K. JILL KIECOLT • TIMUR KURAN • GARY LAFREE • DAVID LAITIN • LINDA D. MOLM • DANIEL MONTI • JEYLAN T. MORTIMER • DANIEL OLSON • DUDLEY POSTON • CECILIA RIDGEWAY • CATHERINE ROSS • RUBÉN G. RUMBAUT • JOACHIM J. SAVELSBERG • DARREN E. SHERKAT • SHELLY SMITH • C. MATTHEW SNIPP • CHARLOTTE STEEH • GILLIAN STEVENS • ROBIN STRYKER • EDWARD TELLES • YU XIE

**International Consulting Editors** JOHN FOX (Canada) • MATTHIJS KALMIJN (The Netherlands) • NOAH LEWIN-EPSTEIN (Israel) • STEVEN L. RYTINA (Canada) • YUVAL YONAY (Israel)

**Editorial Board** ANDREW D. ABBOTT • GARY S. BECKER • CHARLES E. BIDWELL • DONALD J. BOGUE • TERRY N. CLARK • JAMES S. COLEMAN • JOHN L. COMAROFF • ROGER V. GOULD • WENDY GRISWOLD • PHILIP M. HAUSER • EVELYN KITAGAWA • EDWARD O. LAUMANN • DONALD N. LEVINE • DOUGLAS S. MASSEY • WILLIAM L. PARISH • MOISHE POSTONE • EDWARD SHILS • GEORGE P. STEINMETZ • ROSS M. STOLZENBERG • FRED L. STRODTBECK • GERALD D. SUTTLES • LINDA J. WAITE • WILLIAM JULIUS WILSON • KAZUO YAMAGUCHI

THE AMERICAN JOURNAL OF SOCIOLOGY (ISSN 0002-9602) is published bimonthly in July, September, November, January, March, and May by The University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, Illinois 60637. Subscription rates, U.S.A.: institutions, 1 year \$87.00, individuals, 1 year \$39.00. Student subscription rate, USA, 1 year \$27.00 (photocopy of valid student ID must accompany subscription). ASA and BSA members, 1 year \$34.00. All other countries add \$9.00 for each year's subscription to cover postage. Canadian subscribers add 7% GST. Subscription agent for Japan: Kinokuniya Company, Ltd. Single copy rates: institutions \$13.35, individuals \$6.25. Volumes 1-90 available from Schmidt Periodicals GmbH, Dettendorf, D-83075 Bad Feilnbach, Germany. All volumes are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106; and Kraus Reprint and Periodicals, Route 100, Millwood, New York 10546. Subscriptions are payable in advance. Subscriptions will be entered to start with the first issue to be published after receipt of order. **Postmaster:** Send address changes to *American Journal of Sociology*, The University of Chicago Press, Journals Division, P.O. Box 37005, Chicago, Illinois 60637.

ADVERTISING space in THE AMERICAN JOURNAL OF SOCIOLOGY is available, as is rental of its subscriber list. For information and rates, please contact the advertising sales staff, The University of Chicago Press, Journals Division, 5720 South Woodlawn Avenue, Chicago, Illinois 60637. Advertising and list rental is limited to material of scholarly interest to our subscribers.

*Copying beyond Fair Use* The code on the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made beyond those permitted by Section 107 or 108 of the U.S. Copyright Law provided that copies are made only for personal or internal use or for the personal or internal use of specific clients and provided that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc. Operations Center, 27 Congress St., Salem, Massachusetts 01970. *To request permission for other kinds of copying*, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale, kindly write to Permissions Department, The University of Chicago Press, 5801 S. Ellis Ave., Chicago, Illinois 60637. *If no code appears* on the first page of an article, permission to reprint may be obtained only from the author.

Second-class postage paid at Chicago, Illinois, and at additional mailing offices.

© 1995 by The University of Chicago. All rights reserved.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. (∞)



CUR-405426-9-KP18

AJS



American Journal of Sociology

Volume 100 No. 4 January 1995

CONTENTS

9

- 859 Forms of Capital and Social Structure in Cultural Fields: Examining Bourdieu's Social Topography  
HELMUT K. ANHEIER, JÜRGEN GERHARDS, AND FRANK P. ROMO
- 904 Durkheim's Theories of Deviance and Suicide: A Feminist Reconsideration  
JENNIFER M. LEHMANN
- 931 Models for Describing the Underlying Structure of Sex Segregation  
MARIA CHARLES AND DAVID B. GRUSKY
- 972 Vote Turnout of Nineteenth Amendment Women: The Enduring Effect of Disenfranchisement  
GLENN FIREBAUGH AND KEVIN CHEN
- 997 Class Mobility and Political Preferences: Individual and Contextual Effects  
NAN DIRK DE GRAAF, PAUL NIEUWBEERTA, AND ANTHONY HEATH
- 1028 Delinquency and Disdain: Social Capital and the Control of Right-Wing Extremism among East and West Berlin Youth  
JOHN HAGAN, HANS MERKENS, AND KLAUS BOEHNKE

Book Reviews

- 1053 *The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century* by Kim Voss  
JEFFREY HAYDU
- 1055 *Regulating the Social: The Welfare State and Local Politics in Imperial Germany* by George Steinmetz  
ABRAM DE SWAAN

- 1056 *Governing Capitalist Economies: Performance and Control of Economic Sectors*, edited by J. Rogers Hollingsworth, Philippe C. Schmitter, and Wolfgang Streeck  
PETER A. HALL
- 1058 *Forging Industrial Policy: The United States, Britain, and France in the Railway Age* by Frank Dobbin  
RICHARD RUBINSON
- 1060 *Manufacturing Militance: Workers' Movements in Brazil and South Africa, 1970-1985* by Gay W. Seidman  
KEVIN NEUHOUSER
- 1062 *Looking for God in Brazil: The Progressive Catholic Church in Urban Brazil's Religious Arena* by John Burdick  
DAVID A. SMILDE
- 1064 *Religious Melancholy and Protestant Experience in America* by Julius H. Rubin  
RICHARD W. FLORY
- 1066 *Practicing Virtues: Moral Traditions at Quaker and Military Boarding Schools* by Kim Hays  
FRED KNISS
- 1068 *Inside Culture: Art and Class in the American Home* by David Halle  
EUGENE HALTON
- 1069 *Claims to Fame: Celebrity in Contemporary America* by Joshua Gamson  
MICHÈLE LAMONT
- 1071 *Nightwork: Sexuality, Pleasure, and Corporate Masculinity in a Tokyo Hostess Club* by Anne Allison  
YUKO OGASAWARA
- 1073 *Femininities, Masculinities, Sexualities: Freud and Beyond* by Nancy J. Chodorow  
PATRICIA TICINETO CLOUGH
- 1074 *Fallen Women, Problem Girls: Unmarried Mothers and the Professionalization of Social Work, 1890-1945* by Regina G. Kunzel  
KRISTIN KAY BARKER
- 1077 *Birth Weight and Economic Growth: Women's Living Standards in the Industrializing West* by W. Peter Ward  
JEREMY ATTACK

301

Ann 35

K-P-18

- 1079 *Contraception and Abortion in Nineteenth-Century America* by  
Janet Farrell Brodie  
CARL N. DEGLER
- 1080 *Inventing the Feeble Mind: A History of Mental Retardation in  
the United States* by James W. Trent, Jr.  
COLIN SAMSON
- 1082 *Deciding Who Lives: Fateful Choices in the Intensive-Care Nurs-  
ery* by Renee R. Anspach  
CHARLES BOSK
- 1084 *What Machines Can't Do: Politics and Technology in the Indus-  
trial Enterprise* by Robert J. Thomas  
STEVEN PETER VALLAS
- 1086 *Japanese Auto Transplants in the Heart  
Community* by Robert Perrucci  
MICHAEL SCHWARTZ
- 1089 *Defending Community: The Struggle for Development  
in Cedar-Riverside* by Randy Stoecker  
ALFORD A. YOUNG, JR.
- 1091 *Contemporary Urban Sociology* by William G. Flanagan  
HARVEY M. CHOLDIN
- 1093 *Homeless in Paradise: A Map of the Terrain* by Rob Rosenthal  
DANIEL M. CRESS
- 1094 *Constructing White-Collar Crime: Rationalities, Communication,  
Power* by Joachim Savelsberg, with Peter Bruhl  
JAMES WILLIAM COLEMAN
- 1096 *Poor Discipline: Parole and the Social Control of the Underclass,  
1890-1990* by Jonathan Simon  
JOHN R. SUTTON



## IN THIS ISSUE

HELMUT K. ANHEIER is associate professor of sociology at Rutgers University and codirector of the Johns Hopkins Comparative Nonprofit Sector Project. His research interests include structural analysis, comparative sociology, and organizational studies.

JÜRGEN GERHARDS is professor of sociology at the University of Leipzig. His current research includes studies of social movements, the public sphere, and culture.

FRANK P. ROMO is associate professor of sociology at State University of New York, Stony Brook.

JENNIFER M. LEHMANN is assistant professor of sociology at the University of Nebraska. Her research is in social theory, particularly French structuralism and critical theory, including Marxism, feminism, multiculturalism, and cultural studies. She has just published two books on Durkheim.

MARIA CHARLES is assistant professor of sociology at the University of California, San Diego. She received her Ph.D. in 1990 from Stanford University, and results from her dissertation—a cross-national analysis of occupational sex segregation—were subsequently published in the *American Sociological Review*. She is currently involved in various research projects that explore the patterns and processes of occupational attainment and gender stratification in advanced industrial societies.

DAVID B. GRUSKY is associate professor of sociology at Stanford University and coeditor of the Westview Press Social Inequality series. His current research examines the structure of cross-national variability in occupational mobility and sex segregation, the evolution of modern attitudes toward sex-based inequality, and current trends in patterns of occupational and geographic mobility. He is editor of the recently published *Social Stratification: Class, Race, and Gender in Sociological Perspective* and coeditor with James Baron and Donald Treiman of the forthcoming book, *Social Differentiation and Social Inequality: Essays in Honor of John Pock*.

GLENN FIREBAUGH is professor of sociology at Pennsylvania State University and senior scientist at the university's Population Research Institute. He studies social change in the Third World, in post-Communist societies, and in the United States. Recent projects include "Does Economic Growth Benefit the Masses?" (with Frank Beck, *American Sociological Review* [October 1994]). Current work includes a survey on democratization and market values in Romania, and a book manuscript entitled "Analyzing Repeated Surveys."

KEVIN CHEN is a sociologist working as a research associate in the psychiatry department of Columbia University and a student in the Columbia School of Public Health. His research interests include methods, social psychology, medical sociology, health policy, and social change. His

*Political Alienation and Voting Turnout in the United States* was published in 1992. Current research projects include an epidemiological study of drug abuse among American adolescents and youth and a comparative study of health care reform in the United States and China.

NAN DIRK DE GRAAF is research fellow of the Royal Netherlands Academy of Arts and Sciences and is currently affiliated with the University of Nijmegen. He has also conducted research at the Max Planck Institute in Berlin and Nuffield College, Oxford. His main research project concerns the political and cultural consequences of social mobility and mixed marriages. He is a member of the Interuniversity Center for Social Science Theory and Methodology (ICS).

PAUL NIEUWBEERTA is a researcher of the Netherlands' Organization for Scientific Research and is affiliated with the department of sociology at the University of Nijmegen. His main research projects concern the political consequences of social stratification and mobility in Western industrialized countries. His current research focuses on social stratification processes in Eastern Europe.

ANTHONY HEATH is an official fellow of Nuffield College, Oxford, and is joint director of the Centre for Research into Elections and Social Trends. His books include *How Britain Votes* (1985), *Understanding Political Change* (1991), and *Labor's Last Chance?* (1994). He was elected a fellow of the British Academy in 1992.

JOHN HAGAN recently was named W. Grant Dahlstrom Distinguished Professor of Sociology and adjunct professor of law at the University of North Carolina at Chapel Hill. He is coauthor (with Fiona Kay) of *Gender in Practice: A Study of Lawyers' Lives*, coeditor (with Ruth Peterson) of *Crime and Inequality*, and editor of the *Annual Review of Sociology*. He is continuing his collaborative research on the transition to adulthood in Berlin and on lawyers and crime in North America.

HANS MERKENS is professor of empirical educational sciences at the Free University of Berlin. His research interests include educational problems among ethnic minorities, value systems of youth in Eastern and Central Europe, and enculturation processes in corporations.

KLAUS BOEHNKE is professor of socialization research and methodology in the department of sociology of the University of Technology Chemnitz-Zwickau. He earned his Ph.D. in psychology in 1985 at the University of Technology in West Berlin. His main research interests lie in interdisciplinary youth research and in the effects value orientations have for individual psychosocial well-being.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

## Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures will be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract, on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or

money order for \$15.00, payable to The University of Chicago Press in U.S. currency or its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).
2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).
3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).
4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).
5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).
6. Format for references to machine-readable data files is:

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 1/93)

# Forms of Capital and Social Structure in Cultural Fields: Examining Bourdieu's Social Topography<sup>1</sup>

Helmut K. Anheier  
*Rutgers University*

Jürgen Gerhards  
*Wissenschaftszentrum Berlin*

Frank P. Romo  
*State University of New York, Stony Brook*

This article tests one key assumption of Bourdieu's theory of culture fields: that actors are positioned in a "topography" of social relations according to their endowments of economic, social, and cultural capital. Blockmodeling procedures are used to analyze data on German writers and to identify a social structure in which positions vary according to the types and amounts of capital accumulated. A strong split between elite and marginal writers dominates the social structure, and even the fundamental distinction between high and low culture is embedded in this bipartition. Significant differences in both cultural and social capital distinguish elite from nonelite positions; within this bipartition, pronounced differences in cultural capital separate high and low culture. Relative to cultural and social capital, economic capital plays a lesser role in understanding the social structure of cultural fields.

## INTRODUCTION

Recent advances in the sociology of art and literature focus on the structure rather than the content of cultural fields and look at the organizational system in which art is produced, distributed, and consumed (Becker 1982; Blau 1989; Bourdieu 1985a, 1992; DiMaggio 1987; De-Nooy 1992; Peterson 1985). To varying degrees, current work in this realm addresses the long-standing sociological problem of the relation

<sup>1</sup> We would like to thank Friedhelm Neidhardt for his support in conducting this study. Direct correspondence to Helmut Anheier, Department of Sociology, Rutgers University, Lucy Stone Hall, New Brunswick, New Jersey 08903.



between the economic and the cultural and between the material and the artistic. Common to most of these efforts are refinements of Marxian distinctions between structure and superstructure, or Weberian concepts of status and class position, in primarily cultural systems. Usually, such distinctions are conceptualized as dichotomies that classify art, artists, and artistic systems into high and low, serious and light, legitimate and alternative, high brow and low brow, elite and peripheral, or artistic and commercial segments.

Bourdieu (1985a, 1992) has presented the most elaborate theoretical statement about the structure of cultural fields. Bourdieu goes beyond the Marxist concept of class as a system of property rights and introduces a more complex class notion that takes account of different forms of capital—that is, social and cultural as well as economic. In a way, his “field theory,” based on the distinction of class and status (Bourdieu 1985a), “value spheres,” and “spheres of life conduct” (Weber 1972, p. 536), owes as much to Marx as to Weber. Bourdieu positions actors in a social space, or topography, according to economic, social, and cultural characteristics. In contrast to Weber’s macrosociological approach, however, Bourdieu’s field theory is primarily located at the mesolevel. Similar to usage in organizational sociology (DiMaggio and Powell 1983), fields encompass the relations among the totality of relevant individual and organizational actors in functionally differentiated parts of society, such as education, health, and politics, or, as in the case examined here, in art and literature.

Within cultural fields, as in all others, actors are assumed to compete for social positions. This competition gives rise to social structure, which, understood here as a social topology, positions actors relative to each other according to the overall amounts and relative combinations of capital available to them (Bourdieu 1989; Müller 1985, p. 164). The topology is “so constructed that agents who occupy similar or neighboring positions are placed in similar conditions” (Bourdieu 1989, p. 17), which in turn makes such actors more likely to develop similar dispositions, interests, and habits. In analytic terms, Bourdieu defines the structure of a field as a “network, or a configuration, of objective relations” among positions (Bourdieu and Wacquant 1992, p. 97).

Besides its general plausibility, however, Bourdieu’s theory rests on two basic propositions that need critical empirical support: first, the theory assumes that differences in capital endowments are in fact related to the social topography (social structure) of cultural fields in a significant and meaningful way; and second—in a manner similar to DiMaggio’s (1987) classification principles—the theory posits a close correspondence between the social topography and the cognitive structure of cultural

fields, that is, the mental maps actors have of social positions and their organizational and cultural correlates. In this article, we will primarily explore the empirical foundation for the first assumption.

Bourdieu's social topography of fields as a configuration of relations among positions bears considerable affinity to the concept of structural equivalence (Lorraine and White 1973; Sailer 1978), which provides the theoretical bedrock for much recent work in network analysis, specifically blockmodeling (White, Boorman, and Breiger 1976; Wasserman and Galaskiewicz 1993). Structural equivalence is a mathematical principle that groups actors according to their relations with others. Identity in such relations to third parties places actors in equivalent structural positions. By implication, structurally equivalent actors are substitutable, and the set of relations among actors can be reduced to a simpler structure or, mathematically, to a homomorphic image, in which individual actors are replaced by classes of actors, or blocks.

As a formal principle, structural equivalence operationalizes a significant part of Bourdieu's conception of social topography as the relative social positioning of actors in terms of similarity and dissimilarity of social relations. Structural equivalence does not, however, extend to include similarities in interests, cognition, and behaviors, or habitus, as Bourdieu's social topography does (Bourdieu 1985*b*; 1990, chap. 3). Nonetheless, representations of social structure in terms of structural equivalence appear as the first step necessary to examine the extent of similarities in the conditions, interests, and behaviors among actors. Indeed, Bourdieu and Wacquant (1992, p. 105) suggest that we "map out the objective structure of relations" among positions before analyzing the habits and dispositions of actors in a given field. It is, therefore, surprising that despite the clear affinity between the concept of structural equivalence in network analysis and Bourdieu's social topography in his field theory, no systematic empirical tests of Bourdieu's theorizing have been carried out using a relational approach.

We conduct such a test in this article by studying a group of writers and literati in the city of Cologne, Germany. Through the use of social network and categorical data collected with the help of personally administered questionnaires, we look at the pattern of social relations among writers and examine the amounts and types of capital accumulated. Thus, we intend to explore the usefulness of Bourdieu's social topography in understanding the relation between forms of capital and social structure in the field of literature. Specifically, we provide initial tests of Bourdieu's field theory and examine the relative impact of factors such as market position, reputation, and organizational influence on the social positioning of writers.

FORMS OF CAPITAL AND SOCIAL STRUCTURE

Bourdieu's concept of "capital" is broader than the monetary notion of capital in economics; capital is a generalized "resource" that can assume monetary and nonmonetary as well as tangible and intangible forms. Bourdieu (1986, p. 243) distinguishes between three general types of capital, which may assume field-specific contents:<sup>2</sup>

*Economic capital* refers to monetary income as well as other financial resources and assets and finds its institutional expression in property rights. For example, writers differ in the extent to which they earn income from publishing, public readings, or other literary activities.

*Cultural capital* exists in various forms. It includes long-standing dispositions and habits acquired in the socialization process, the accumulation of valued cultural objects such as paintings, and formal educational qualifications and training. For example, through the study of literature and fine arts, writers may acquire tastes and styles distinct from others. Important for our purposes is the distinction between *incorporated* cultural capital, in the form of education and knowledge, and *symbolic* cultural capital, the capacity to define and legitimize cultural, moral, and artistic values, standards, and styles. High-culture genres and writers of literary criticism may have high degrees of symbolic capital, whereas writers in other genres, such as folk art, may enjoy little.

*Social capital* is the sum of the actual and potential resources that can be mobilized through membership in social networks of actors and organizations. For example, writers may differ in the size and span of their social networks to cultural institutions.

The types of capital differ in liquidity and convertibility and in their potential for loss through attrition and inflation. Economic capital is the most liquid, most readily convertible form for transformation into social and cultural capital. By comparison, the convertibility of social capital into economic capital is costlier and more contingent; social capital is less liquid, "stickier," and subject to attrition. While it is difficult to convert social into cultural capital, the transformation of cultural into social capital is easier.

The differences in the liquidity, convertibility, and loss potential of forms of capital entail different scenarios for actors in social fields. Some positions are characterized by high volumes of economic capital, yet

<sup>2</sup> Bourdieu's usage and definitions of the various forms of capital is sometimes somewhat cursory. His paper (Bourdieu 1983) on forms of capital, available in English in 1986, offers the most complete and systematic discussion of capital forms; see also Müller (1985), Gartman (1992), and Bourdieu and Wacquant (1992).

lower volumes of cultural and social capital; others will rank high in terms of cultural capital, yet somewhat lower in other forms. The *nouveaux riches*, for example, are typically well endowed with economic capital relative to a paucity of cultural capital. International business consultants rely on high levels of social capital, relative to cultural and economic capital, and intellectuals typically accumulate higher amounts of cultural and symbolic capital than they do economic and social endowments. Writers of avant-garde literature may have little economic capital available to them, but they may appear well endowed in terms of symbolic-cultural capital; in contrast, writers of romance novels may rank high in accumulated economic capital, yet low in cultural capital as well as in relevant social capital.

### Dominant Forms of Capital and Social Structure

Bourdieu (1985a; 1992, pp. 165–81) argues that the dominance of specific forms of capital is characteristic of different types of social fields. Of particular importance for cultural fields is the distinction between the field of *restricted cultural production* and the field of *large-scale cultural production*. Both fields differ to the extent to which economic and non-economic capital forms become dominant. The field of restricted production is relatively autonomous from market considerations. Economic success is secondary to symbolic value, and writers compete for cultural capital in the form of recognition, reputation, and legitimacy rather than for monetary rewards. In contrast, the field of large-scale cultural production is characterized by the predominance of economic considerations and market success. In the large-scale case, writers compete as producers in a market to seek financial returns first and foremost.

Based on this reasoning, the social topography of cultural fields appears bipartite (Bourdieu 1992). The two segments, however, seem largely unrelated: in one segment, writers are artists; in the other they are commercial producers. The distinction between the fields of restricted and large-scale production suggests rigid boundaries imposed upon segments. Each field forms a segment in which different primary “currencies”—prestige versus money—are exchanged; and each segment is internally structured according to either market or reputational success. Different mechanisms operate within each segment. In the field of restricted production, the distinction between the writing of text and its evaluation and legitimation as “literature” introduces hierarchical elements into the social structure. Critics and “metaliterati” act as the cultural legitimator and fabricate “creative interpretation for the benefit of the creator” (Bourdieu 1985a, p. 18). Critics provide data for the

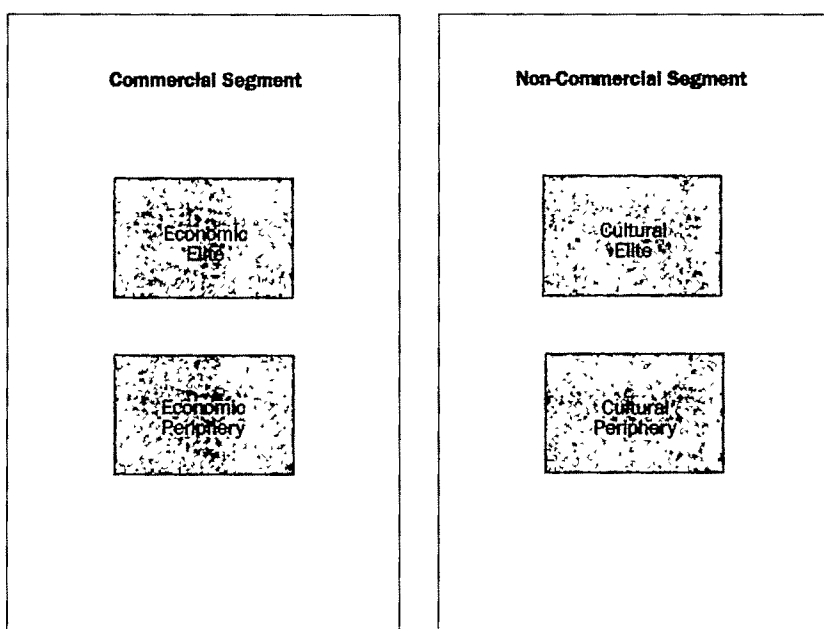


FIG. 1.—Basic model of restricted and general cultural production

competitive audience of peers as the writers' alter egos and offer the basic material for the reputational hierarchy in the struggle for symbolic, cultural, and social capital.

Thus the basic model of the structure of cultural fields includes two dimensions: first, the distinction between economic and noneconomic capital in the constitution of two separate segments of literary production; and second, the internal stratification of each dimension into dominant and peripheral groups of writers. Bourdieu's thinking would lead us to expect a social topography with two segments and four groups, as graphically represented in figure 1.<sup>3</sup>

The distinction between the fields of restricted and large-scale production of culture is ideal-typical; and any cultural product is, as a manifestation of economic and noneconomic capital, both commodity and symbol (Bourdieu 1989). If we move beyond the dichotomy of economic versus noneconomic capital, different combinations of capital may involve intermediary positions, and Bourdieu (1985a) acknowledges that cultural fields may contain a variety of intermediate forms between art as a com-

<sup>3</sup> We would like to thank the *AJS* reviewer who suggested this expected structural outcome of Bourdieu's topography.

mercial commodity and as a symbolic and cultural good. For these intermediate positions, commercial success and critical acclaim do not necessarily coincide. Social capital, for example, may be a more important means of status competition in high-culture genres than in mass culture; moreover, some writers may regard membership in professional organizations and circles as inopportune, perhaps even contrary to their self-understanding as artists, while others see it as an important ingredient for the accumulation of the economic, cultural, and social capital needed for status competitions.

Although not fully developed in his writing, Bourdieu's (1979, 1983, 1985a) work suggests that the dominance of different forms of capital may correspond to distinct configurations in the social topography of cultural fields. This implies that the basic model presented in figure 1 is not the only one possible and that different topologies of social positions may result from different capital forms. It is useful to engage in a "thought experiment" to explore the relations between forms of capital and social structure. We do so by relating the dominant form of capital to two types of partitions in social structures: segmentation and hierarchy.

*Segmentation* refers to the number of relatively distinct, structurally separate, and unrelated parallel components of the social structure. Typical of segmentation in culture are the basic distinctions between restricted and large-scale production of cultural goods or the symbolic differences between "high culture" and "low culture," "serious literature" and "light literature." A segmentation is represented in the partition between the commercial and the noncommercial segments in figure 1.

*Hierarchy* refers to the extent to which partitions yield clusters of social positions in terms of status differences. Typical hierarchical elements of cultural fields include the positions of prominent writers as the elite and the unknown, "struggling" writers as the periphery. Segmentation elements are relationally independent, but as social positions are linked across different statuses hierarchies emerge. In figure 1, hierarchies are found in the elite-periphery split within each segment.

In reference to DiMaggio (1987, pp. 447-52), who explores the relationship between industry structure and artistic classification systems, we can construct three types of *ceteris paribus* relations between dominant forms of capital (as the independent variable) and the partition of social structure (as the dependent variable) in the field of modern literary production. Each form of capital adds specific tendencies to the degree of segmentation and hierarchy in the social structure.

A predominance of economic capital leads to weaker degrees of segmentation between genres that are based on market success. The hierarchy of the social structure would be pronounced, though relatively fluid,

to reflect graduated distinctions in terms of economic success, as commercial producers try to reduce the strength of genre distinctions in search of larger audiences as markets.

A predominance of social capital leads to multiple though weakly institutionalized segments that are maintained, along the lines of differentiated genres, by professional organizations and other loci of interest mediation that often take the form of complex social networks. Hierarchies are less pronounced than they would be under the predominance of economic capital, a flexibility that reflects membership participation and ranks in formal and informal associations.

A predominance of cultural (symbolic) capital tends to lead to highly segmented and hierarchical social structures. The social structure would be divided into a segment of legitimized art, or high culture, and another segment of nonlegitimate art, or low culture. The relationship between the two segments is hierarchical, as are relationships between the position of legitimator and producer of literature within the high-culture segment. Moreover, the high-culture segment is stratified according to reputational status positions, as competition among writers leads them to innovate and develop new styles.

Thus economic, social, and cultural capital differ in terms of the structural patterns they generate. Economic capital differs from social capital in terms of hierarchy and from cultural capital in terms of segmentation; social capital differs from cultural capital in terms of hierarchy. Table 1 contains a representation of the major components of an operational model of forms of capital and social structure in art. Each of the three forms of capital implies a basic distinction. Economic capital focuses on commercial success versus failure with money as the major currency and economic status as the major indicator. Social capital distinguishes between membership and nonmembership in professional organizations and informal networks with "contacts" as the major currency and differences in membership affiliation as the major indicator. Cultural capital marks a distinction between recognition and indifference in the perception and reception of literature with prestige as the major currency and reputation and education as indicators; symbolic cultural capital makes a distinction between art and nonart or high and low culture with artistic legitimation as the major currency and genre hierarchies as the major indicators.

In each case, we hypothesize specific forms of partitions in terms of segmentation and hierarchy. Economic capital operating alone will result in a social structure characterized by low segmentation and strong hierarchies. Social capital by itself leads to high segmentation and weak hierarchies, and cultural capital produces strongly segmented and hierarchical social structures.

TABLE 1  
AN OPERATIONAL MODEL OF FORMS OF CAPITAL AND SOCIAL STRUCTURE IN CULTURAL FIELDS

Form of Capital	Basic Distinction	Major Currency	Degree of Segmentation	Degree of Hierarchy	Indicators
Economic .....	Monetary success versus failure	Money	Weak	Strong	Economic status
Social .....	Member versus nonmember	Social contacts and connections	Strong	Weak	Membership
Cultural .. .....	Recognition versus indifference	Prestige	Strong	Strong	Reputation, education
Symbolic cultural	Art versus nonart	Legitimation	Strong	Strong	Genre hierarchies



In blockmodel images, a segmentary social structure between two blocks, A and B, is indicated by

	A	B
A	1	0
B	0	1,

and, in a weaker form, as

	A	B
A	1	0
B	0	0,

where neither of the two segments (blocks) relates to the other; instead, each segment constitutes a distinct, separate component. In other words, there are relations among the positions of each block (indicated by "1"), but none between blocks (indicated by the "0" off-diagonal entries in the two image matrices). As mentioned above, the partition between restricted and general production is an example of segmentation in cultural fields.

Next, hierarchical social structures are indicated by nonzero off-diagonal entries in the image matrices. A strong hierarchy is represented by the image matrix

	A	B
A	1	0
B	1	0,

where block B relates to A, but not vice versa. Positions in A are related, whereas positions in B are not. The relationship between elite writers (block A) and a disparate, unrelated set of unknown writers trying to gain access to elite positions (block B) would be an example of a strong hierarchy. A weak hierarchy,

	A	B
A	1	0
B	1	1,

ranks two interrelated blocks, for example, the relation between a group of established elite writers (block A) and a cohesive circle of younger

writers (block B) aspiring to seek the friendship or advice of the first group. The center-periphery model,

	A	B
A	1	1
B	1	0,

is a variation of the simple hierarchy, whereby block B represents a set of positions that, while unrelated to one another, have some form of mutual relationship with the interconnected positions of block A. For example, a cohesive group of writers occupying central positions in the social structure relate to peripheral writers who are themselves unrelated to each other.

Of course, actual social structures are composites of such basic segmentary and hierarchical elements. The positions of elites in the social topography of writers are of particular interest. Do different forms of capital exchanged within the same cultural field lead to a single elite or to multiple elites? Do elites differ in the extent to which they represent financial achievements and reputational success or in their ability to legitimize and to organize their interests? Is it true, as Mannheim (1940, pp. 82–83) suggests, that we can detect in the field of culture organizing elites distinct from artistic or political elites? Can we find partitions similar to those of fragmented elites in politics and intellectual circles in general (see Kadushin 1974, 1976), with counterelites leading the rebellion against established art and its representatives? Do elites form an inclusive network linking all top positions in the social structure, as suggested by the integrated elite model? How do peripheral writers compare to elite writers in terms of capital and composition of forms?

The following analysis proceeds in two steps. First, we describe the social structure and attempt to identify social positions in terms of segmentation and hierarchy. Second, we examine how social positions differ according to the overall amount and composition of capital.

## DATA AND METHODOLOGY

Our research design and data collection focused on the social networks among writers in the city of Cologne, Germany.<sup>4</sup> Rather than taking a sample, we decided to include the total population of writers living in

<sup>4</sup> Art and literature in Germany are not dominated by a single cultural capital; the country is characterized by several competing cultural centers such as Cologne, Munich, Hamburg, and Berlin.

the greater metropolitan area. We defined *writer* as any producer of fictional texts, thereby excluding authors of science, travel, and "how-to" literature. We applied neither aesthetic, artistic, social, nor any other criteria to differentiate between prominent and unknown writers, between "serious" and "light" literature, between high culture and popular culture. We used several published and unpublished directories in addition to gathering information from publishers, critics, cultural institutions, and local writers' groups to identify 222 writers.<sup>5</sup> We managed to conduct personal interviews with 150 (67.6%) of the 222 writers by administering a semistandardized questionnaire.<sup>6</sup>

The social network among writers was measured in four dimensions, or types of ties, using an "aided-recall" method by presenting subjects with the complete listing of the 222 writers living in or near Cologne; these one-page lists were submitted to respondents in the course of asking the various network questions during the interview.

1. Familiarity with the work of other writers (AWARENESS):<sup>7</sup> "On this list, would you please check the names of those authors whose work is familiar to you?"
2. Friendship ties to other writers (FRIENDSHIP):<sup>8</sup> "On this list, would you please check the names of those authors whom you consider as friends?"
3. Received help and assistance from other writers (ASSISTANCE; results are derived from a Boolean union of the following two questions):<sup>9</sup> "On

<sup>5</sup> In addition to published directories of writers, we consulted lists of local writers maintained at municipal and other libraries, the membership lists of the local chapter of the German Writers' Association (VS), and numerous conversations with individual writers themselves.

<sup>6</sup> Writers were contacted by letter and phone, appointments were made, and in most cases interviews took place at the writer's residence. We conducted some interviews and some were conducted by students of the University of Cologne. The interviews lasted about 90 minutes. To the extent possible we collected data on the missing cases. Using a number of secondary sources such as recent editions of *Kurschners Literaturlexikon* (the most complete directory of German writers available), we succeeded in gathering data on age, sex, and number of publications. For all three variables, we found no statistically significant differences between the 150 writers we interviewed and the missing cases.

<sup>7</sup> This is a cognitive tie measuring awareness of other writers, similar to "citation" ties in other network studies (see Burt 1982; Burt and Minor 1983; Romo and Anheier 1991).

<sup>8</sup> Friendship and other close personal relations have often been identified as crucial factors in understanding the background and history of writers and their literary work, e.g., the friendship cult in the late Enlightenment, literary clubs, salons, and coffee houses (Back and Polisar 1983; Gerhards 1986).

<sup>9</sup> This included discussion of manuscripts, assistance in finding a publisher, and help in gaining access to cultural institutions, e.g., help to initiate and establish contacts to arrange public readings

this list, would you please check the names of those authors with whom you have discussed manuscripts in the past?" "On this list, would you please check the names of those authors who were helpful in establishing contacts with publishers?"

4. Loyalty and reference ties (INVITATION):<sup>10</sup> "On this list, would you please check the names of those writers you would like to invite for dinner?"

Based on these data, we constructed four binary matrices, whose  $ij$ th entry is "1" if a specific tie exists between writer  $i$  and writer  $j$ . Since possible distortions introduced by missing cases are difficult to assess in relational data structures, we excluded from the analysis writers who had not been interviewed and several writers who declined to answer the questions listed above. This reduced the number of valid cases to 139. A blockmodel of the 139 writers and four types of ties (White et al. 1976) was obtained by applying the ICON-H algorithm, an operationalization of the structural equivalence principle and a hierarchical clustering procedure based on combinatorial optimization (Romo and Anheier 1991, 1994).

With multiple binary incidence matrices as input, ICON-H seeks to partition the population into  $c > 1$  blocks "to separate as effectively as possible high-density from low-density regions" (Boorman and Levitt 1983, p. 2) in the permuted and blocked matrices. By utilizing an iterative hill-climbing method in maximizing a target function (ratio of between-block sum of squares to within-block sum of squares), ICON-H calculates a local optimum.<sup>11</sup> Repeated applications of ICON-H yield a target function (goodness of fit), which seems to suggest a social structure with seven blocks; in this case, partitions beyond the seven-block level yield singletons, in which individual writers occupy identified social positions. The appendix offers a brief discussion of the mathematical, theoretical, and technical background on ICON-H and compares ICON-H to alternative blockmodel techniques.

Table 2 presents descriptive statistics for the seven-block solution and the results of a principal component analysis to obtain reliability and

<sup>10</sup> DiMaggio (1986) uses a similar measurement in his study of American resident theaters.

<sup>11</sup> An alternative blockmodel technique uses Johnson's (1967) diameter (maximum) method, a clustering algorithm based on Euclidian distances, as implemented in the STRUCTURE program (Burt 1982, 1987). As reported in Anheier and Gerhards (1991), it yields similar results and interpretations of the social structure, as did the hierarchical cluster analysis (CONCOR) reported in Gerhards and Anheier (1989), although the ICON-H solution is clearly superior to both distance-based and factor-analytic models. See the appendix for a fuller discussion of the methodological background to ICON-H and the overall goodness of fit of the blocking solution in relation to alternative algorithms.

TABLE 2  
DESCRIPTIVE STATISTICS

Block	Description	Block Size	Size* (%)	Variation† (%)	Average Prominence‡
A .....	Cultural elite	6	4.3	49.04	.752
B .....	Organizational elite	5	3.6	73.46	.531
C ..	Subelite	20	14.4	86.01	.311
D .....	Semiperiphery 1	22	15.8	79.81	.217
E .....	Semiperiphery 2	33	23.7	86.03	.078
F .....	Light culture	10	7.2	91.35	.022
G .....	Periphery	43	30.9	87.45	.033
Total .....		139	100.0		

\* This is the relative block size as a percentage of the total number of writers in the network.

† Percentage of variance in the distance among occupants of position *B*, (block), which is accounted for by a single principal component.

‡ Aggregate prominence ranges between 1 and 0; 1 = the most prominent person in the network, 0 = the least prominent person(s). Aggregate prominence reported here is the mean aggregate prominence for each block.

goodness-of-fit indicators (Burt 1982, 1987) that are useful for an initial independent validation of the ICON-H results. Except for block A, all blocks show acceptable proportions of variance explained by a single component in the distance among occupants of a position (block). Moreover, except in the case of the elite, reliability indicators are at least .9 or better for all but 10 of the 108 writers in the semiperiphery and the periphery. Thus, reliability indicators suggest an acceptable blockmodel solution (see the appendix for a comparative assessment of the ICON-H solution). Table 3 presents the density matrices for the four types of ties, whereas table 4 shows the image matrices.

## RESULTS AND DISCUSSION

To identify different elements in the social structure in terms of segmentation and hierarchy, we interpret the mutual (relative) absence of interblock relations as an indication of segmentation, while the presence of unequal interblock relations signifies hierarchy. For example, the virtual absence of relations between block G (periphery) and block F (light culture) in all four matrices in table 3 would suggest a segmentation. Moreover, the rows and columns for the peripheral block G for all four types of ties show an absence, rather than a presence, of relations both internally and externally to other blocks. The unequal relationship between block E (semiperiphery 2) and block A (cultural elite) would be an example of a hierarchical element in the social structure: as shown in the first

TABLE 3  
DENSITY MATRICES FOR SEVEN-BLOCK SOLUTION\*

	G	F	E	D	C	B	A
AWARENESS							
G .....	1	3	1	2	5	5	13
F .....	2	64	2	6	4	2	7
E .....	3	5	6	6	19	24	60
D .....	5	5	9	21	39	68	84
C .....	2	1	2	5	15	26	43
B .....	10	6	17	53	77	95	90
A .....	3	0	7	20	31	60	70
FRIENDSHIP							
G .....	0	0	0	0	1	1	0
F .....	0	8	0	1	1	0	0
E .....	0	1	2	1	0	1	2
D .....	0	2	1	3	3	8	9
C .....	0	0	1	1	6	9	8
B .....	0	0	1	6	5	10	17
A .....	0	0	2	8	4	17	20
ASSISTANCE							
G .....	0	0	0	0	1	0	2
F .....	0	11	1	1	0	0	0
E .....	1	1	2	2	1	5	8
D .....	0	1	1	3	2	11	16
C .....	0	0	1	1	1	7	9
B .....	0	0	1	5	2	5	20
A .....	0	2	2	8	5	20	27
DINNER INVITATION							
G .....	0	0	0	1	1	1	4
F .....	0	8	0	0	1	0	3
E .....	0	1	1	1	4	2	14
D .....	0	0	1	4	5	10	14
C .....	0	0	1	1	3	4	10
B .....	0	0	1	12	14	40	57
A .....	1	0	4	11	17	20	37

NOTE —For descriptions of blocks A–G, see table 1.

\* Densities—the ratio between possible and measured ties—range between 0 (no ties present) and 100 (all possible ties present). Mean densities for each matrix are AWARENESS = 22.5; FRIENDSHIP = 3.3; ASSISTANCE = 3.8, DINNER INVITATION = 6.3

TABLE 4  
IMAGE MATRICES FOR SEVEN-BLOCK SOLUTION\*

	G	F	E	D	C	B	A
AWARENESS							
G .....	0	0	0	0	0	0	0
F .....	0	1	0	0	0	0	0
E .....	0	0	0	0	0	1	1
D .....	0	0	0	0	1	1	1
C .....	0	0	0	0	0	1	1
B .....	0	0	1	1	1	1	1
A .....	0	0	0	0	1	1	1
FRIENDSHIP							
G ... ..	0	0	0	0	0	0	0
F .....	0	1	0	0	0	0	0
E ... ..	0	0	0	0	0	0	0
D ... ..	0	0	0	0	0	1	1
C .....	0	0	0	0	1	1	1
B ... ..	0	0	0	1	1	1	1
A .....	0	0	0	1	1	1	1
ASSISTANCE							
G . ....	0	0	0	0	0	0	0
F .....	0	1	0	0	0	0	0
E . ....	0	0	0	0	0	1	1
D .....	0	0	0	0	0	1	1
C . ....	0	0	0	0	0	1	1
B .....	0	0	0	1	0	1	1
A . ....	0	0	0	1	1	1	1
DINNER INVITATION							
G ....	0	0	0	0	0	0	0
F ....	0	1	0	0	0	0	1
E ....	0	0	0	0	0	0	1
D .....	0	0	0	0	0	1	1
C .....	0	0	0	0	0	0	1
B .....	0	0	0	1	1	1	1
A .....	0	0	0	1	1	1	1

NOTE — For descriptions of blocks A–G, see table 1  
 \* Cutoff densities are based on mean density densities below and equal to the mean density = 0; densities above the mean = 1 Cutoff densities for each matrix are AWARENESS > 22.5, FRIENDSHIP > 3.3; ASSISTANCE > 3.8, DINNER INVITATION > 6.3

part of table 3, the density of AWARENESS from block E to A is 60%, but only 7% from A to E.

The blockmodel analysis reveals a general partition of the social structure into two major segments (fig. 2).<sup>12</sup> The first segment constitutes the periphery of the social structure (blocks F and G) and consists of 38.1% of the writers. The second segment includes all other blocks (A–E), and represents the core segment, with 61.9% of the writers. As we will see, however, the segments are very different. While the core segment is internally differentiated in terms of a hierarchical social structure, the peripheral segment is not, and instead of hierarchical elements, we find a further segmentation, embedded in the first, that separates the light-culture block from a larger group of peripheral writers.

Very few ties link the core and peripheral segments. In table 3, AWARENESS has a density of 13%, the highest relation between core and periphery; most other interblock densities range between 2% and 5%. Within the peripheral segment, the independence between the set of peripheral writers in block G and the light-culture block F is even more pronounced. Interblock ties are absent in the matrices FRIENDSHIP, ASSISTANCE, and DINNER and, with 2% and 3%, are extremely low for AWARENESS. The light-culture block is a segment contained within the structure of a larger segment of peripheral writers, which seems to indicate a split in the social structure embedded within the larger and dominant pattern that sets the core segment apart from the peripheral segment. Thus we find the strong segmentary pattern,

$$\begin{array}{cc} 1 & 0 \\ 0 & 1, \end{array}$$

replicated in the relation between the core and light culture, and the weak segmentation,

$$\begin{array}{cc} 1 & 0 \\ 0 & 0, \end{array}$$

between the core and the periphery.

By contrast to the segmentary patterns, the five blocks in the core segment reveal pronounced hierarchical elements, either in terms of strong and weak hierarchies or in terms of center-periphery configurations. This segment comprises two elite blocks made up of 7.9% of all the writers in the study, a semielite representing 14.4% of the sample,

<sup>12</sup> The labeling of the different strata and blocks will become evident as we continue presenting our results. The labels are based primarily on the structural location of blocks while taking into account average prominence scores of block members.



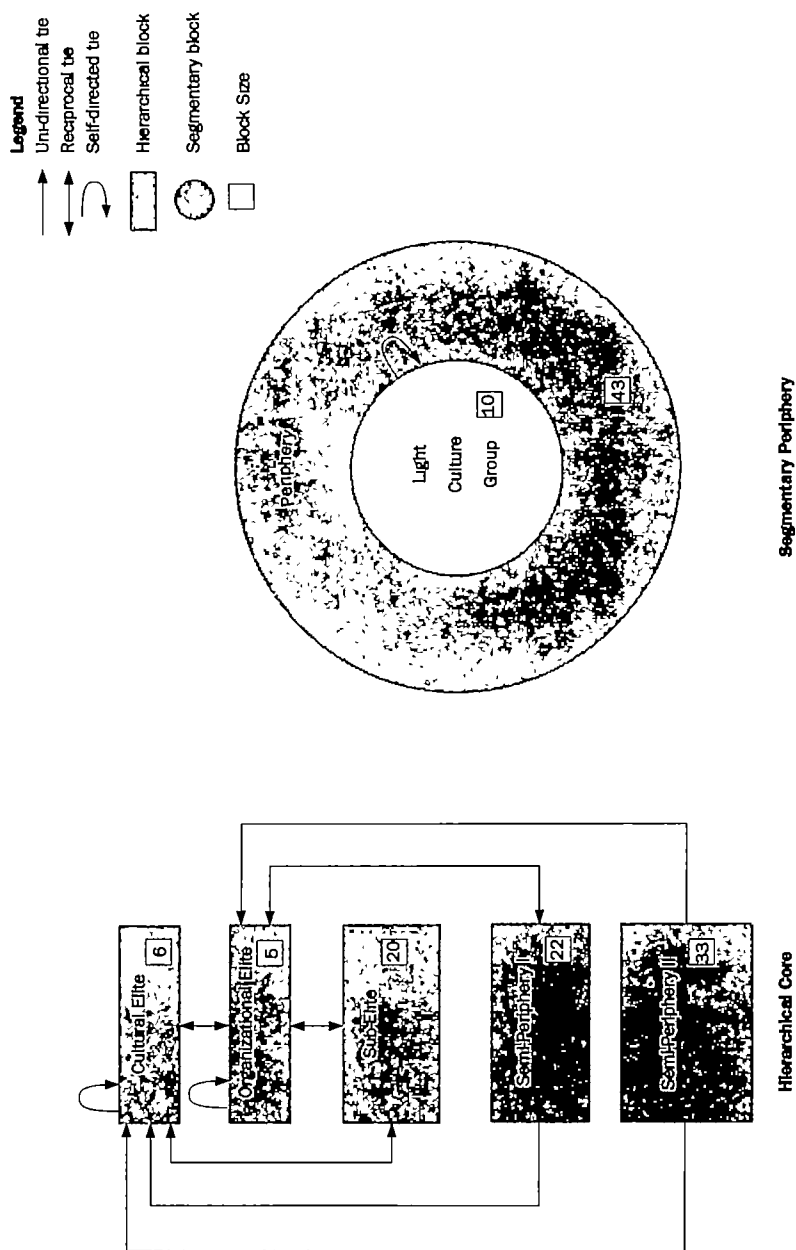


FIG. 2.—Simplified graph of block relations

and two larger, semiperipheral blocks with 30.2% of the sample (see tables 1 and 2). In this respect, it is useful to observe the prominence scores displayed in table 2.<sup>13</sup> Average prominence declines from .752 and .531 for the elite blocks within the core segment, to .311 for the subelite and .218 for semiperiphery 1 and .078 for semiperiphery 2. For the periphery, the average aggregate prominence score declines to .033 and to .022 for the light-culture block embedded within it.

Although we will examine block relations more closely further below, this first glance at the overall block structure seems to suggest that the primary distinction in cultural fields may not be segmentation into high culture and low culture, as predicted by Bourdieu (1983, 1985a). Instead, as reflected in figure 2, we find a pronounced segmentation between a hierarchical core and a largely undifferentiated peripheral segment, with the distinction between high and low culture embedded in the periphery. We will now take a closer look at some blocks.

*Elite.*—The elite group is divided into a cultural elite of six writers, an organizational elite of five, and a larger block, the subelite, with 20 writers.<sup>14</sup> Members of the cultural elite in block A combine high degrees of cultural capital and prestige with the gatekeeper role of an elite based on privileged access to central institutions. As we will show below, the cultural elite form the aesthetic *and* reputational elite primarily, and an organizational elite only secondarily. In contrast, members of the organizational elite, block B, are still relatively prestigious writers with an average prominence score of .531; however, compared to the cultural elite, they are less known for their cultural, aesthetic, and literary significance as writers than for their centrality in organizational matters.

Several examples may illustrate the differences between the cultural and the organizational elite:<sup>15</sup> the cultural elite group includes well-known literary critics and editors of prestigious literary series and programs, such as the editor of a highly influential literature journal who is acclaimed not only for his own literary work but also for having discovered and supported "young literary talents." Another member of the cultural elite is perhaps the most widely known author among the respondents; one of his latest novels found international recognition and was,

<sup>13</sup> Aggregate prominence, a measure of individual network structure (Knoke and Burt 1983, pp. 206–7; Burt 1982, 1987) assigns "1" to the most prominent individual in a network. All less prominent network members have scores expressed as fractions of the maximum prominence, with "0" indicating the least prominent member in the network.

<sup>14</sup> We provide a narrative for elite blocks only; the larger semiperipheral and peripheral blocks are described with the help of correspondence analysis below.

<sup>15</sup> The following descriptions are based on information we collected about elite writers in addition to the data gathered in the course of the personal interviews.

for example, reviewed in the *New York Times Book Review* and other gatekeeping publications in cultural centers. He is not very active in professional organizations, with the exception of PEN (International Association of Poets, Playwrights, Editors, Essayists, and Novelists). Because of his high reputation, he is invited to virtually all major cultural and literary events in the greater Cologne area. Previously editor-in-chief at a leading publishing house, he is credited for his support of younger writers.

Consider, by contrast, a typical profile of a member of the organizational elite. One of the most recognized members of block B is known as "organizer" and "literature manager"; he is very active in the German Writers' Association (VS), the formal representative body of German writers; he organizes public readings and is involved in most larger organizational events. Despite his prominence as a writer, he is primarily known for his qualities as an organizer. Another member of the organizational elite is the chief editor of the literature program section of Germany's largest radio and TV station. He is also very active in professional associations such as the VS and has been involved in the publishing of a prominent literary magazine.

Both cultural and organizational elite members are connected with leading publishing houses. It seems difficult to establish a strict rank order among publishers of elite writers: Fischer, Suhrkamp, Reclam, and Rowolth are high-volume, high-quality publishers, followed by Insel, Hoffmann and Campe, and Kiepenheuer and Witsch. Claasen, Gutenberg, and Schneekluth are smaller yet well-regarded literature presses. In contrast to peripheral writers, publishing houses of elite authors are household names.

*Subelite.*—Not only because of its larger size but also because of its very structural position, it is more difficult to present an illustration of the subelite (block C), which includes a greater variety of literary types and genres than the two elite blocks. Members of this block are still relatively well known, although their average aggregate prominence amounts to only 41% of that of the cultural elite and to 59% of that of the organizational elite. In contrast to elite writers, members of the subelite are typically not full-time writers; they do, however, tend to work in other, related cultural areas. The subelite of block C includes a well-known journalist and documentary writer who only recently shifted to fine art genres, a recent recipient of the best foreign novel prize from the Academie Française, a lesser-known Berthold Brecht student who is a theater critic, and a high-ranking official from city hall who writes novels in his spare time, as well as representatives of the young literary scene, including a police officer and author of poetry, drama, short stories, literary essays, and children's stories.

## Relational Patterns

How and to what extent do the various blocks relate to each other? Figure 2 presents a simplified graph of the information contained in the density matrices of table 3, and the image matrices in table 4.<sup>16</sup> The structural relations between blocks emphasize why we have chosen the labels "elite," "semiperiphery," and "periphery." Observe that, in all four matrices of table 3, block densities are lowest in the upper-left-hand corner and highest in the lower-right-hand corner, a pattern that indicates a more complex relational structure among elites than among peripheral positions. Similarly, in the image matrices of table 4, "1"s cluster in the lower-right-hand corners, whereas virtually all other entries are zero. In figure 2, note how core blocks are hierarchically arranged, while blocks in the peripheral segment are not.

*Periphery.*—Relations within and among peripheral blocks are virtually absent, and writers in these blocks tend to be unaware of each other's work. A density of "0" for FRIENDSHIP means that among the 43 writers of the periphery no friendship ties exist; neither do assistance relations or dinner invitations. In fact, zero is the most frequent density of relations between the periphery and all other blocks in the network. Since the absence of interblock relations serves as an indication of segmentation in the social structure, peripheral writers form a separate segment of positions next to writers in the core of the social structure. Thus we conclude that the relationship between core and periphery is segmentary, not hierarchical.

*Light culture.*—The light-culture group is embedded in the periphery and does not constitute an independent block as such (see the discussion of embedded segmentation in the appendix). Structurally, this group is the result of an internal segmentation within the peripheral segment of a social structure dominated by a core-periphery split. Members of this group are essentially folk artists and poets writing verse in the local vernacular, a genre and language regarded neither as fine art nor high German. Members of this block are part of the local culture rooted in medieval Catholic traditions, and they are not included in the national literary culture. While literary works written in some of the other German dialects may be acceptable among literary critics and publishers, the Cologne vernacular does not normally qualify as an artistic medium. Relationally, the folk art group, in the German context regarded as low

<sup>16</sup> Image matrices represent a summary picture of the information contained in the density matrices and are often used for both analytical and descriptive purposes in blockmodel analysis. Entries at and below the mean density are recorded as "0," and entries above the mean density as "1." Image matrix representations of blockmodel results bring the essential features of the social structure into focus.

or light culture, represents an "isolated island" in the social structure and forms a segment of its own based on genre distinction. Internal AWARENESS, FRIENDSHIP, ASSISTANCE, and DINNER INVITATION ties are the highest of all the groups, with the exception of the two elite blocks.

*Semiperiphery.*—The members of the periphery and semiperiphery 1 and 2, together representing about two-thirds of all writers, have in common the relative absence of internal block relations. What differentiates periphery and semiperiphery, however, is the pronounced elite orientation of the latter. Moreover, these three groups differ somewhat in the extent to which ties sent to the elite blocks are reciprocated: while all blocks seem to be unaware of the periphery, with only very few friendship and assistance ties or dinner invitations, the two blocks in the semiperiphery receive the recognition of the elite's awareness to at least a small extent, with densities of 17% and 53% from the organizational elite and 7% and 20% for the cultural elite (see table 2).

In many respects, the relational pattern of the semiperiphery is located between that of elite and periphery. In terms of internal awareness of each other's work, the major diagonal in the first section of table 2 indicates that the intrablock density of semiperiphery 1 is 21% higher than that of semiperiphery 2 (6%) and the periphery (1%). Both semiperiphery 1 and 2 are essentially elite oriented, and relations to nonelite blocks are either absent or weak. In turn, semiperipheries receive more attention from the elite than they do from peripheral blocks. The relationship between elite blocks and the semiperipheral blocks can be characterized as either a strong hierarchy, that is,

$$\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0, \end{array}$$

as in the case for the cultural elite block on the one hand and for both semiperiphery 1 and 2 on the other; or as a center-periphery configuration, that is,

$$\begin{array}{cc} 1 & 1 \\ 1 & 0, \end{array}$$

for the relationship between the organizational elite and the semiperiphery 2 block.

*Elite and subelite.*—Turning to the elite blocks in the core of the social structure, we find that relational patterns indicate a hierarchy among elite positions. The cultural elite block shows high internal densities

throughout (70%, 20%, 27%, and 37%) whereas the organizational elite block reveals high intrablock densities in the case of AWARENESS (95%) and DINNER INVITATIONS (40%) only (table 3). In contrast, the subelite block shows low internal densities for all types of ties, with 15% for AWARENESS, 6% for FRIENDSHIP, 1% for ASSISTANCE, and 3% for DINNER INVITATIONS. In terms of ties sent, the subelite group shows a pattern similar to that of semiperiphery 1 and 2; it differs in the pattern of ties received. Here, the subelite group receives higher density ties from both the cultural elite and the organizational elite. Overall, however, the result is a center-periphery configuration with the two elite blocks as the center and the subelite in the position of the periphery.

The organizational elite group shows higher densities of awareness ties to other blocks than does the cultural elite group. In terms of awareness ties received, however, the situation is reversed: the cultural elite group receives a higher density of ties. This pattern is to some extent also present in other relations, whereby the cultural elite group receives higher density ties from nonelite blocks throughout, including the subelite. Moreover, the two elite blocks differ in the way they relate to each other. Except for AWARENESS, densities of ties sent from the organizational elite to the cultural elite are higher than intrablock densities. For the cultural elite, the pattern is reversed: ties to the organizational elite are less pronounced than internal relation for all four types of ties measured. This pattern of unequal relations between blocks seems to suggest a weak hierarchy between the cultural elite and the organizational elite.

*Summary.*—The blockmodel analysis identified two segments: core and periphery. The first segmentation between core and periphery dominates the social structure, whereas the second segmentation is embedded in the peripheral segment and differentiates the light-culture block and the larger group of peripheral writers. Only the core segment revealed a hierarchical structure, with elite, subelite, and semiperiphery 1 and 2 as major components. While elements of strong and weak hierarchy were present, particularly between the semiperiphery and the cultural elite, the overall network pattern can be interpreted as a core-periphery structure with the following key elements:

1. Elite blocks form the center and relate primarily to themselves and, to a lesser extent, to the semiperiphery.
2. The organizational elite and the cultural elite form a weak hierarchy that is dominated by the cultural elite.
3. The relations between the elites and the subelite reveal a clear hierarchy in terms of a center-periphery configuration.
4. Semiperipheral blocks tend to relate to the elite blocks as the center, rather than to themselves, and are part of a hierarchical system dominated by the elite blocks.

5. The periphery appears structurally isolated, both internally as well as in relation to other blocks, and forms a segment rather than a hierarchical component of the cultural field.
6. The embedded segmentary structure of the light-culture block is a high-density area located like an "island" in the low-density periphery of the social structure.

How does the identified social structure relate to Bourdieu's social topography? First, the structure differs from the basic model suggested by Bourdieu's thinking and represented in figure 1. The identified structure is more complex, particularly in respect to the position of the semi-peripheries and the elite differentiations. Second, the basic partition is not between general and restricted production of culture or between high and low culture; rather the dominant partition suggests a core-periphery split as the basic organizing principle. The blockmodel analysis shows a pattern indicating segmentation *and* hierarchy, whereby parts of the social structure are primarily segmentary (the periphery and the light-culture block) and others are primarily hierarchical (i.e., the core segment with elites, subelite, and semiperipheries). The identified social structure does not reveal the relational properties of low segmentation *and* weak hierarchy that would be expected if the structure were based on economic capital alone. Moreover, it does not correspond to the expected outcome of social capital, which by itself would lead to high segmentation and weaker hierarchies. In fact, the identified properties are closest to those expected under the dominance of cultural capital, which will produce a strongly segmented and hierarchical social structure.

### Correspondence Analysis

Having described the social structure of the cultural field in terms of hierarchy and segmentation, we now examine how identified social positions (blocks) relate to different amounts and forms of capital endowments. We do this with the help of correspondence analysis (Greenacre 1984; Weller and Romney 1990; Benzecri 1992)—a multivariate technique extensively employed by Bourdieu (1979) in *Distinction*—which enables us to describe and analyze the relationships between variables such as social position and capital endowment. Similar to principal component analysis, it projects rows and columns into a lower-dimensioned vector space. Unlike principal component analysis, which is used for data representing continuous measurements, correspondence analysis accepts nominal-, ordinal-, and interval-level data. This makes correspondence analysis particularly suitable for many social science data problems in which data are typically qualitative and of lower-level measurement.

The data are taken from our survey of writers. The blocks (positions) form the columns and the various categorical measures of capital forms constitute the rows of a contingency table, which provides input to our correspondence analysis and contains the frequencies for the number of times block membership (columns) coincides with particular attributes of categorical variables. For example, the input data would list the number of times members of the elite block were recipients of literary honors and would also contain the same information for all other blocks. Note that correspondence analysis, like related methods such as principal component analysis and multidimensional scaling, neither establishes nor requires any notion of causality between variables—in our case, between block positions, capital forms, and amounts. What the analysis allows us to do, however, is to see to what extent different social positions “correspond” to different forms and amounts of capital.

To examine Bourdieu's (1989, p. 17) argument that positions in a social topography differ by the overall volume and by the composition of types of capital, we select different sets of indicators for each capital form (see table 1 above).

In the case of economic capital, and the underlying distinction between commercial success and failure, we expect positions to differ according to economic variables such as overall income,<sup>17</sup> the proportion of income earned through literary activities,<sup>18</sup> and if writers can normally make a living from literary income alone.<sup>19</sup> For example, elites would be expected to derive a higher proportion of income from literary and cultural activities and thus be more likely than peripheral writers to live off literary income.

In the case of social capital,<sup>20</sup> the distinction between inclusion and exclusion and membership in formal professional associations, special interest organizations, and informal, local literary circles and debating

<sup>17</sup> Income was measured as an ordinal variable with seven ranges indicating increasing average monthly incomes. The obtained distribution was split at the median into “high” and “low” incomes and cross-tabulated with the column variable, BLOCK MEMBERSHIP.

<sup>18</sup> Professional income was measured as the average proportion of average monthly net income derived from activities as a writer and artist, such as royalties, honoraria, including income from other cultural activities. The split between low and high proportions of “cultural” income occurred at the median.

<sup>19</sup> This variable is based on the following questions: “Taking everything into account, can you normally make a living from your activities as a writer?” Answers were coded yes or no and cross-tabulated with the variable block membership to yield input data for the correspondence analysis.

<sup>20</sup> We selected measures independent of the network questions we used for the blockmodel analysis above. For this reason, we did not include sociometric measures such as reachability or span that can be derived from the same data.



clubs differentiate position in the social topography. For example, elite writers are more likely to be members of core interest organizations such as VS. We also include a measure for the extent to which subjects have friendship contacts in cultural fields other than their connections to other writers included in our study.<sup>21</sup>

Unequal possession of cultural capital is part of the differentiation between the elite and the periphery (Bourdieu 1985a; DiMaggio 1987). Familiarity with literary history, styles, and tastes and the ability to converse in the language of literary criticism are largely acquired with a university education. In the absence of formal professional socialization, a university-level education in liberal arts and literature (in this case German literature) becomes essential cultural capital and credentials for elite access. Thus we include a variable indicating if a writer studied literature at a university. Elites are in large part the reputational nobility of the social structure. Thus we would expect that elite writers would be members of prestigious invitation-only national and international literary societies such as PEN.<sup>22</sup> As PEN is primarily a reputational association and not an interest association, membership in the club indicates cultural instead of social capital. In contrast to peripheral writers, elites are expected to have received more honors and more official recognition in the form of grants and stipends.<sup>23</sup>

For symbolic cultural capital, the genre distinctions between high culture and low or light culture (e.g., literature in high German vs. the local vernacular), as well as the distinction between the producers and critics of fictional texts, are expected to differentiate among segments and hierarchical positions.<sup>24</sup>

Three coefficients are important in interpreting the results of correspondence analysis (Greenacre 1984, 1990). Inertia (INR), which ranges between zero and one, indicates to what extent rows (capital forms and amounts) and columns (block membership) determine the model and its axes. Inertia coefficients sum to unity for rows and columns, respectively.

<sup>21</sup> We asked a sequence of questions about the number of friends in various fields and institutions of culture and art, such as music, painting and sculpture, theater, radio and television, and museums. The combined distribution of "cultural friends" was split at the median into "few" and "many" cultural contacts.

<sup>22</sup> This variable is based on the question: "Are you a member of PEN?" and answers are coded in dichotomous form.

<sup>23</sup> The variable is based on the Boolean union of answers to two questions: "Have you ever received any literary prizes or official recognitions?" and "Have you ever received a stipend or grant for your literary work?" The resulting joint distribution was split at "1," indicating that the writer has won a prize, stipend, or grant.

<sup>24</sup> The two variables are based on the following questions: "Do you write in the local vernacular (*Mundartliteratur*)?" and "Do you write literary essays?"

For example, in table 5 we see that, for the column variables (block membership), "periphery" determines the entire model to 10% ( $INR = .100$ ), the first axis to 9.3%, the second to 4.6%. For the row variables in table 6, we observe that, among the social capital variables, membership in formal associations determines the model to 11.2%, the first axis to 16.7%, and the second to 2.2%. In other words, inertia coefficients tell us how important a form of capital is for particular social positions and, vice versa, how important a social position is for a particular capital form. The squared correlations (COR) range between zero and one and express the proportion of variance in the rows and columns that are explained by the axes.<sup>25</sup> In table 6, we see, for example, that under cultural capital the first axis accounts for 81.5% ( $COR = .815$ ) of the variation of the attribute, "Member PEN: no." Thus, the squared correlations indicate how well the axes capture the variations in amounts and forms of capital on the one hand and variations in social positions on the other. Location (LOC) refers to the coordinates of the positions (blocks) and categorical variables (amount and forms of capital) in the lower-dimensional vector space, and ranges between plus and minus one. If we imagine a two-dimensional coordination system, we locate the column variable "periphery" at coordinates .155 on the first axis, and at .069 on the second axis (table 5). The row variable "formal membership: high" is situated in the opposite quadrant at coordinates  $-.572$  and  $-.130$ . As in multidimensional scaling, the correspondence between columns (block membership) and rows is indicated by the proximity of their locations in the lower-dimensional vector space.

The results of the correspondence analysis indicate a four-factor (or axis) model.<sup>26</sup> However, the four factors carry different weights (see table 7): the first is dominant and carries 59.90% of the total inertia, followed by 23.60% for the second, and three additional axes with negligible inertias of 8.57% for the third, 6.91% for the fourth, and 1% for the fifth. For purposes of this analysis, we will primarily focus on the first two axes, which together represent almost 84% of the total inertia in the model. To aid interpretation of the detailed results presented in table 6, figure 3 offers a graphical representation of the first two axes and the location of social positions and selected categorical variables in this two-dimensional space.

<sup>25</sup> In interpreting the results of the correspondence analysis, we employ the conventional cutoff criterion of .35 for squared correlations.

<sup>26</sup> For computational reasons, because of the small block sizes, we had to collapse the organizational and culture elites (blocks A and B) to form a joint elite block. For this reason, we cannot follow up on some of the important differences that exist between the two elite blocks in the context of the correspondence analysis.

TABLE 5  
COLUMN CONTRIBUTIONS

Block	TOTAL MODEL		FIRST AXIS			SECOND AXIS			THIRD AXIS			FOURTH AXIS		
			COR		INR	LOC		COR	LOC		COR	LOC		COR
	COR	INR	LOC	COR	INR	LOC	COR	INR	LOC	COR	INR	LOC	COR	INR
Periphery . . . . .	.991	100	.155	.559	.093	.069	.108	.046	-.030	.021	.024	.115	.304	.440
Light culture . . . . .	1.000	.283	.469	.411	.194	-.557	.581	.697	.060	.007	.022	-.029	.002	.006
Semiperiphery 2 . . . . .	.999	.120	.143	.303	.061	.160	.374	.190	.111	.181	.253	-.098	.141	.244
Semiperiphery 1 . . . . .	.998	.075	-.093	.137	.017	-.019	.005	.002	-.199	.629	.554	-.120	.227	.247
Subelite . . . . .	.957	.139	-.349	.919	.213	-.030	.007	.004	-.005	0	0	.064	.031	.062
Elite . . . . .	.989	.282	-.655	.894	.422	-.156	.051	.061	.146	.044	.146	-.002	0	0

NOTE.—INR is the degree to which block membership/capital form determines model and its axis (dimensions); inertia ranges between 0 and 1. COR is the proportion of variance in block membership and capital form explained by the axis, range is between 0 and 1. LOC is the location of a variable in the lower dimensioned vector space, range is -1 and +1

A look at the various forms and indicators of capital (table 6) shows that the three economic capital variables determine the model to 11.4% (total inertia of .005, .009, .015, .035, .013, .037), social capital to 30.7%, cultural capital to 57.9%, which includes symbolic cultural capital with 29.3% of total inertia. This indicates that, with 88.6% of total inertia, noneconomic forms of capital are dominant in the positioning of writers in the social topography of the literary field. Moreover, three indicators of social and cultural capital, that is, membership in PEN (14.2%), membership in formal associations (20.8%), and the genre distinction between literary criticism and other forms of writing (19%), together determine the model to over 50%. Each of these three indicators alone carries more explanatory power than the three economic capital variables combined.

As seen from the social positions (table 5) the first and most important axis establishes the difference between elite, subelite, and semiperiphery 1 on one side, and semiperiphery 2, light culture, and periphery on the other (fig. 3). Here, the elite blocks and semiperiphery 1 load negatively, while all other blocks load positively. The second axis establishes the distinction between popular culture and high culture—in particular it sets the popular culture block apart from semiperiphery 1. The following two axes relate to the difference between semiperiphery 2 and the periphery. Thus, the first axis indicates the differences in forms of capital between the hierarchical core and the segmentary periphery, the second axis between light and high culture, and the third and fourth relate the differences among peripheral and semiperipheral positions. The first axis, however, is by far the most important—with a proportion of the total inertia of about 60%—and shows that the distinction between core and peripheral positions runs across most indicators of capital forms. It is only within this distinction that the difference between high and light culture is played out, as the second axis demonstrates.

Figure 3 illustrates that core and periphery are almost mirror images of each other. Not all variables, however, are equally important in pointing out differences in types of capital held by elite and nonelite positions. Significantly, one cultural capital variable (membership in PEN) and one social capital variable (formal membership in professional associations) determine the first axis to almost 50% with a combined inertia of .497.

*Economic capital.*—Elites earn a higher proportion of average monthly income from cultural activities; they are also more likely to be able to make a living from literary incomes. Peripheral writers earn lower income proportions from being a writer and are less likely to make a living from writing. Elite and peripheral positions do not, however, differ in the amount of income available to them; as the results indicate, they differ in the structure of economic capital.

*Social capital.*—Elite writers are more likely to be members of formal

TABLE 6  
ROW CONTRIBUTIONS

VARIABLE	TOTAL MODEL		FIRST AXIS			SECOND AXIS			THIRD AXIS			FOURTH AXIS		
	COR	INR	LOC	COR	INR	LOC	COR	INR	LOC	COR	INR	LOC	COR	INR
Economic capital:														
Income: low .....	.999	.005	.016	.017	0	.011	.009	0	-.054	.203	.012	-.105	.769	.056
Income: high .....	.707	.009	-.039	.044	.001	.062	.111	.004	.072	.147	.016	.119	.405	.055
Proportion income: low .....	.980	.015	.148	.670	.016	.030	.028	.002	-.092	.258	.044	.028	.024	.005
Proportion income: high ..	.941	.035	-.344	.815	.048	-.110	.084	.012	.068	.032	.013	-.039	.010	.005
Makes a living from writing: no .....	1.000	.013	.113	.477	.011	-.002	0	0	-.077	.221	.034	.090	.301	.058
Makes a living from writing: yes ..	.989	.037	-.377	.667	.041	.035	.006	.001	.169	.135	.058	-.197	.181	.097
Social capital:														
Formal membership: low ..	.995	.096	.477	.893	.143	.103	.041	.017	.090	.032	.036	.085	.029	.040
Formal membership: high ..	.997	.112	-.572	.893	.167	-.130	.046	.022	-.097	.028	.034	-.108	.032	.052
Informal membership: low ..	.999	.014	-.095	.341	.008	-.081	.247	.015	-.090	.307	.051	.052	.104	.021
Informal membership: high ..	.996	.059	.383	.358	.035	.293	.208	.052	.368	.329	.226	-.205	.102	.087
Culture friends: low .....	1.000	.007	.034	.083	.001	-.042	.124	.004	.019	.025	.002	-.103	.767	.079
Culture friends: high .....	.993	.019	-.078	.055	.002	.113	.118	.009	-.031	.009	.002	.298	.812	.222

Cultural capital:													
University literature													
study no	1.000	.030	.192	.467	.023	-.157	.309	.039	.111	.155	.054	.074	.069 .030
University literature													
study: yes	.997	.033	-.223	.458	.026	.187	.320	.045	-.122	.137	.053	-.094	.082 .039
Member PEN: no	.988	.018	.157	.815	.025	.070	.161	.012	-.012	.005	.001	-.015	.007 .002
Member PEN: yes	.978	.124	-1.000	.779	.162	-.516	.180	.095	.137	.013	.019	.092	.006 .010
Honors received: low	1.000	.027	.219	.841	.037	.056	.055	.006	-.025	.011	.003	-.073	.093 .036
Honors received: high	.997	.054	-.451	.812	.074	-.127	.065	.015	.073	.021	.013	.151	.091 .071
Symbolic cultural capital													
Literary essays: no	.997	.056	.320	.582	.054	-.175	.175	.041	-.196	.219	.142	-.059	.020 .016
Literary essays: yes	.999	.047	-.266	.554	.043	.147	.170	.034	.181	.257	.140	.048	.018 .012
Dialect literature: no	.998	.026	-.106	.250	.011	.181	.732	.081	-.025	.014	.004	.009	.002 .001
Dialect literature: yes	1.000	.164	.664	.266	.073	-1.000	.710	.493	.188	.021	.041	-.066	.003 .006

NOTE — INR is the degree to which block membership/capital form determines the model and its axis (dimensions), inertia ranges between 0 and 1 COR is the proportion of variance in block membership and capital form explained by the axis, range is between 0 and 1 LOC is the location of a variable in the lower dimensional vector space, range is between -1 and +1.

TABLE 7

AXES, INERTIAS, AND PERCENTAGES OF INERTIA

Axes	Inertia	%
1 .....	.0805	59.90
2 .....	.0317	23.60
3 .....	.0115	8.57
4 .....	.0093	6.91
5 .....	.0014	1.02
Total ...	.1344	100.00

NOTE — $\chi^2 = 202.15$ ,  $df = 105$ 

professional associations and less likely to be part of informal literary circles and clubs. The number of friends in the field of culture does not differentiate between elite and peripheral positions. The results indicate that elite and peripheral writers do not differ in the same way across various measures of social capital. What seems to matter are formal professional structures. Being a "joiner" of literary circles and clubs as such seems indicative of peripheral positions.

*Cultural capital.*—We find that elite writers are more likely to have studied literature in a university setting, to be members of PEN, and to have received prizes and honors. Thus, elite writers possess significantly higher amounts of all three forms of cultural capital than peripheral writers.

*Symbolic cultural capital.*—Differences in symbolic cultural capital, as expressed in the division of labor between the critic and the producer of literature, are forcefully replicated in the distinction between elite and peripheral groups. Elite writers are not only producers but also judges of literature. By being writer and critic at the same time, elite members are in a position to evaluate not only their peers but also to review the periphery. Through its metaliterary activity, the elite can offer cultural legitimacy to peripheral and semiperipheral writers.

The second axis of figure 3 separates the low culture segment from all other blocks, exemplified in the difference between the light culture block and semiperiphery 1. The variable "dialect literature" determines the second axis to over 50% with an inertia of .574. The two positions do not differ significantly in terms of economic and social capital. The only other difference is found in the area of cultural capital, with light culture writers being less likely to have studied literature at the university level.

The third axis sheds more light on the semiperipheral positions. Results suggest that semiperiphery 1 and 2 differ primarily in one form of social

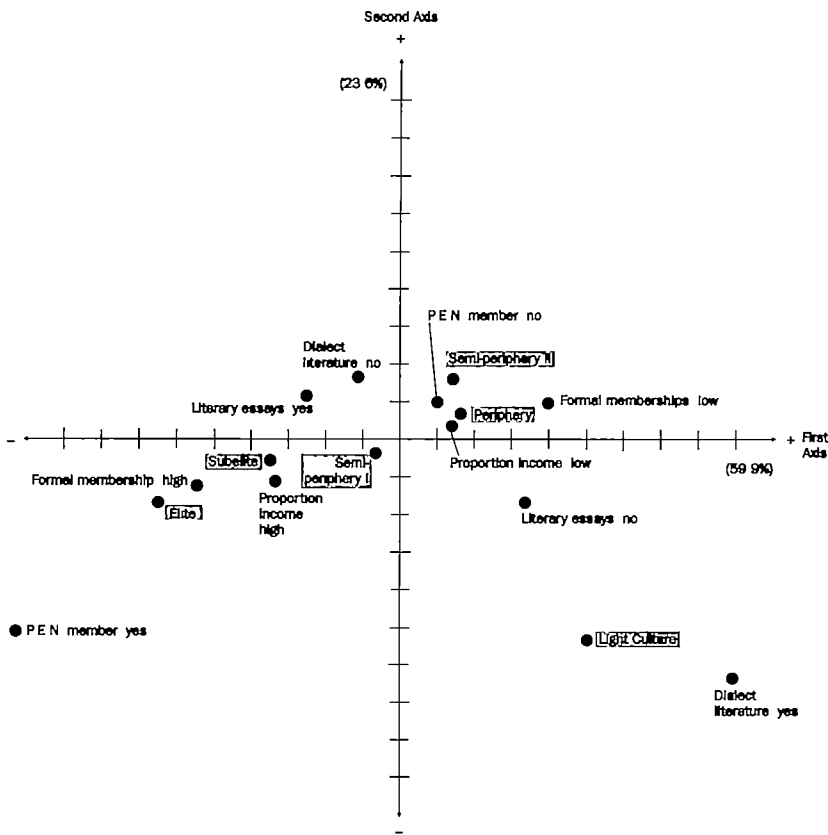


FIG. 3.—Graphical representation of correspondence analysis

capital: the slightly less peripheral writers in semiperiphery 1 tend to rank higher in terms of membership in informal literary circles and clubs. The fourth axis, finally, highlights differences between periphery and semiperipheries. The inertia and squared correlation suggest that these groups differ primarily according to social capital (number of cultural friends) and the amount of economic capital.

In summary, we find that significant differences in cultural and social capital distinguish elite from nonelite positions, pronounced deficiencies in symbolic cultural capital characterize the low culture segment embedded in the periphery, differences in social capital help account for the two semiperipheral positions, and lower economic and social capital distinguish periphery from semiperiphery. Moreover, just as is the case for the dominant hierarchical split in the social structure in the blockmodel analysis, the primary axis here does not differentiate between high culture



and light culture, as one would expect from Bourdieu's theory, but between core and periphery. Only the second, and less important, axis differentiates high from low culture. Significantly, this second differentiation occurs in the periphery itself and is subsequent to the primary distinction between core and periphery. Thus the results of the correspondence analysis fully support what we learned from the blockmodel analysis above.

## CONCLUSION

Bourdieu has made significant contributions to sociological class analysis. By developing theoretical links to Weber's distinction between class and status, "value spheres" and "life conduct," he was able to go beyond the primacy of economic capital in Marxist conceptions of class and to lay the foundation for a more refined analysis of the social structure of modern societies, specifically cultural fields. Our analysis here examined the relative importance of different forms of capital in the field of literature. We showed how the various forms of capital correspond to a social structure with multiple elites and peripheries—a social topography more complex than the model based on the dichotomy between restricted and general production of cultural goods. In general, however, we found strong support for Bourdieu's (1989, p. 17) hypothesis that actors are distributed in social space by both the overall volume and relative composition of capital.

The blockmodel analysis identified two segments: core and periphery, with the popular-culture segment embedded in the latter. The core segment revealed a hierarchical structure, with elite, subelite, and semiperiphery as major components. By contrast, the periphery was non-hierarchical. Overall, the blockmodel analysis indicated a pattern of segmentation *and* hierarchy, whereby parts of the social structure are more segmentary and less stratified and others are primarily hierarchical—that is, the core segment with elites, subelites, and semiperipheries. We found that the general properties of the social topography are closest to those expected under the dominance of cultural capital.

For the identified social structure among writers, cultural capital proved the dominant factor in the differentiation of social positions. The positions of elites, semielites, semiperipheries, and periphery are closely related to differences in the amounts of accumulated cultural capital. At the same time, we found that other forms of capital are important, too. We can, however, assume that the primary positioning in artistic fields happens through the accumulation of cultural capital. Once amounts of cultural capital sufficient to maintain social status within the field have been accumulated and secured, additional gains may be converted to other forms of

capital. For example, available cultural capital may be beneficial in seeking prominent positions in interest associations (social capital) and may in turn open up new income opportunities (economic capital).

We identified a social structure in which social positions vary according to the types and amounts of capital accumulated. Significant differences in cultural, symbolic, and social capital distinguish elite from nonelite positions, pronounced differences in symbolic cultural capital separate out the low-culture segment, differences in social capital help account for the two semiperipheral positions, and lower economic and social capital distinguish periphery from semiperiphery. Overall, the results suggest a gradual diminishing of noneconomic capital as we move from the hierarchical core of the social structure to its segmentary periphery.

Within the elite, we detected a differentiation between an organizing elite distinct from artistic or political elites as suggested by Mannheim (1940, pp. 82–83). However, elites did not appear fragmented, as has been observed in political fields, nor did we detect counterelites attempting to challenge established art and its representatives. The core-periphery split, and not internal elite cleavages, dominate the social structure. Even the fundamental distinction between high and low culture is embedded in this general bipartition of the social structure examined here. In short, our analysis suggests that, for peripheral writers, it does not really matter what type of literature they produce—high culture, low culture, avant-garde, or pulp fiction—rather what seems to matter primarily is that they are not part of the elite system. As the correspondence analysis indicated, two organizational factors, membership in PEN and in formal writer's associations, are key to the capital deficiencies of the periphery. In large part, therefore, it is the organizational incorporation of cultural capital (PEN membership) and social capital (membership in formal professional associations), and not the distinction between high and low culture or restricted and commercial production of art that seems to serve as the primary force in the social structure of cultural fields.

While we agree with Bourdieu that sociology is a social topography (1989, p. 16), he has not fully exploited the structural implications of this statement in his empirical work. For example, in his most elaborate empirical work, *Distinction*, social positions were neither defined nor measured in relational terms, as suggested by the principle of structural equivalence in social network analysis, but reduced to nominal, sometimes ordinal variables as representations of occupational groups and rankings as proxy measures. Bourdieu's theory, however, would demand a more structural, empirical approach to explore its usefulness. It is, after all, as Gartman (1992) pointed out, perhaps the most complex general theoretical framework to the study of cultural systems since critical theory. While Bourdieu's theory of cultural fields is well grounded in classi-

cal sociological theorizing, and while it has considerable plausibility, it nonetheless rests on two key assumptions: first, it assumes that differences in capital endowments determine the social topography of cultural fields to a significant extent, and second, that the theory posits a close correspondence between the social topography and the cognitive structure of cultural fields. While this analysis provided systematic empirical support for the first assumption, the second still needs to be explored.

## APPENDIX

### Blockmodeling Methods

In network analysis, interpersonal choices, typically obtained during the administration of questionnaires, are used to construct square ( $N \times N$ ) binary adjacency matrices, one matrix for each choice. These individual matrices (i.e., networks) are stacked to create a compound data structure. In the data collected on German writers, 139 answered four such interpersonal choice questions (see text above). The compound network structure, then, has 556 rows and 139 columns. A technique of mathematical taxonomy called blockmodeling provides a versatile way of identifying complex relational patterns contained in such compound network structures (see, e.g., Lorraine and White 1971; Breiger, Boorman, and Arabie 1975; White et al. 1976; Boorman and White 1976; Arabie, Boorman, and Levitt 1978; Boorman and Levitt 1983). The technique is based on Lorraine and White's (1971) theoretical logic of "structural equivalence of objects." This particular organizing principle postulates that social structure is composed of characteristic bundles of interlocking relations among a population of objects (e.g., actors, organizations). These characteristic bundles of relations reveal structure because they consistently link specific sets of objects to other sets of objects in the population. In the strictest sense, objects  $a$  and  $b$ , belonging to the same population  $C$ , are considered structurally equivalent if, for any relation  $R$ , and any other object  $x$  also belonging to population  $C$ :

$$aR, x \Leftrightarrow bR, x,$$

and,

$$xR, a \Leftrightarrow xR, b.$$

In other words, objects  $a$  and  $b$  are structurally equivalent to one another if  $a$  is related to or, in this case, *chooses* any other object  $x$  on  $C$  in exactly the same way  $b$  does, and if any other object  $x$  chooses  $a$  in the same way  $x$  chooses  $b$ . In such a case, objects  $a$  and  $b$  are a maximal homogeneous group. This deterministic definition of structural equivalence

lence is far too rigid to be useful in studying human populations. A complete inventory of all social choices among human actors will be impossible to obtain, even in small organizations. For practical purposes, network analysts are forced to sample types of social choices from members of a population. The activity of sampling relations introduces measurement error and, consequently, this deterministic concept is unrealistic and impossible to apply.

Given imperfect measurement and stochastic noise, we adopt a statistical definition of structural equivalence of objects. That is, we operationalize the structural equivalence principle from a least squares point of view. In Boorman and Levitt's (1983, pp. 1-3) formulation of this method,<sup>27</sup> the central data structure is a multigraph  $\alpha(\mathbf{N}; \{\mathbf{A}_k\}_{k=1}^m)$ , where  $\mathbf{N}$  is a set of  $n$  entities (usually a set of specific persons) and  $\{\mathbf{A}_k\}_{k=1}^m$  is a family of directed graphs contained in  $\mathbf{N} \times \mathbf{N}$ , typically identified with  $(0, 1)$  binary incidence matrices  $[a_{pq}^k]_{m \times n}$  (usually associated with a set of social networks recording various types of social choices and evaluations that people have of one another). The primary objective of blockmodeling is to partition  $\mathbf{N}$  into  $r > 1$  blocks,  $\mathbf{P} = \{b_1, b_2, \dots, b_r\} \in \ell(\mathbf{N})$ ; that is, partition lattice on  $\mathbf{N}$ . In general, such a partition requires the effective separation of the high-density from the low-density regions of  $\mathbf{A}_k$  (i.e., high- and low-density submatrices resulting from permuting and partitioning the binary incidence matrices with the information in  $\mathbf{P}$ ). Here  $|\mathbf{P}| = r$  and reflects the degree of aggregation or blockmodeling refinement ( $1 < r < n$ ) desired by the investigator.

This goal of effective separation can be made concrete by seeking to maximize the following objective target function  $T$  over  $\mathbf{P}\{\mathbf{P} \in \ell(\mathbf{N}) \text{ \& } |\mathbf{P}| = r\}$ :

$$T = \sum_{k=1}^m \left[ \sum_{i=1}^{|b_i|} \sum_{j=1}^{|b_j|} S_{(b_i, b_j)} (D_{ij}^k - \mu_k)^2 \right], \quad (\text{A1})$$

where,

$$S_{(b_i, b_j)} = (|b_i| |b_j|) - (\delta_{ij} |b_i|), \quad (\text{A2})$$

$$D_{ij}^k = \frac{1}{S_{(b_i, b_j)}} \sum_{p \in b_i} \sum_{q \in b_j} a_{pq}^k, \quad \mathbf{A}_k = [a_{pq}^k]_{m \times n} \quad (\text{A3})$$

$$\mu_k = \frac{1}{n(n-1)} \sum_{p, q=1}^n a_{pq}^k. \quad (\text{A4})$$

<sup>27</sup> Early formulations of this method can be found in Kelly (1980) and Romo (1980). For a treatment that explicitly ties this method to the maximization of the between sum of squares and the minimization of the within sum of squares, see Romo (1986).

Here,  $\delta_{ij}$  is Kronecker's delta, and  $|b_i|$  and  $|b_j|$  equal the size of the set indexed by the  $i$ th row block and  $j$ th column block ( $i, j = 1, 2, 3, \dots, r$ ). Finally, note that in the Boorman-Levitt formulation, any contribution from elements on the principal diagonal,  $a_{ii}^k$ , is eliminated: that is,  $a_{ii}^k$  is coded "0" in  $A_k$ .

A FORTRAN algorithm, ICON-H (Iterative Combinatorial Optimization on Networks with Hierarchical extensions), has been developed to maximize  $T$  (i.e., to achieve formally local optima through a stepwise hill-climbing methodology) over multiple binary incidence matrices typically associated with choice networks.

Our analytical interests involve testing whether *aggregates* of writers display consistent cultural characteristics and whether we can observe variation in such characteristics across aggregates, where such aggregates are defined by applying the aforementioned operationalization of structural equivalence (ICON-H) to relational data. This concern for social aggregates is directly applied to our decisions about the level of blockmodeling refinement (i.e., the number of blocks) used to represent structure in relational data. Any blockmodeling algorithm, including ICON-H, is capable of producing blocks with a single member (i.e., unit blocks). In this investigation, we are not interested in such unit blocks. Thus, in applying the ICON-H algorithm to the German writers' networks, we calculated block assignment vectors (i.e., the partition/permutation vector  $\mathbf{P}$ ) starting at the two-block level of refinement ( $r = 2$ ). Then we continued to refine the solution (i.e., increase the number of blocks) until unit blocks began to appear. Here, a terminal point ( $tp$ ) in the process of block refinement is identified as the solution in which the first unit blocks appear. A likely candidate for the preferred level of refinement is the one that is found in the solution obtained just prior to the solution that produces unit blocks. Thus, if one or more unit blocks appear in the solution for the seven-block level of refinement ( $r = 7$ ), then the preferred solution would be the six-block level of refinement ( $r = 6$ ). Given the demands of any particular investigation, coarser solutions (i.e., higher levels of aggregation and a smaller number of blocks) may be desirable. Coarseness is achieved by further rolling back the level of refinement from  $tp$ . Indeed,  $tp - r_p$  (where  $r_p$  equals the number of blocks in the preferred solution) is a convenient measure of the coarseness of  $r_p$ . For the German writers' network,  $tp$  is reached at the seven-block level ( $tp = 7$ ). Thus, we utilize the six-block level ( $r_p = 6$ ) in this analysis.

For every level of refinement tested ( $2 \leq r \leq 8$ ), we started with 100 block assignment vectors ( $\Pi_{bc}$ ,  $r = 1, 2, \dots, 7$ ;  $c = 1, 2, \dots, 100$ ), each representing a random assignment of individual writers to  $r$  blocks. The random initial conditions ( $\Pi_{bc}$ ) was used as an initial condition upon

which the ICON-H algorithm based its search of a locally optimal solution (i.e., the partition/permutation vector  $\mathbf{P}_{bc}$  where the objective target function  $T$  can no longer be increased by rearrangements obtained from the stepwise, hill-climbing methodology of the algorithm). This strategy produces 100 solutions (many reproducing one another). From this set of solutions, we chose the one in which  $T$  is maximum.

### *Embedded Segmentation*

Note that we have just claimed that we use a six-block-level model in this analysis. Yet, in the main body of this paper, we present a seven-block-level model. The discrepancy is a result of the methodology employed to extract the seventh block. The concept of "embedded segmentation" implies that structurally equivalent blocks are hierarchically embedded in the global structure or in parent blocks. Indeed, the light-culture block is such an embedded segment, extracted from the large ( $N = 53$ ) periphery segment. ICON-H permits investigators to hold constant the parent permutation/partition vector (i.e., the parent block assignment vector) and seek hierarchical segments within a given parent block. All information contained in the parent block assignment vector is used by the algorithm to discover embedded segments.

One immediate question is, Why use hierarchical methods to discover additional blocks? Or, to put this question another way, Is it not enough to simply increase the number of blocks ( $r$ ) and directly apply the combinatorial methodology discussed above? The answer, of course, depends upon the nature of the data being analyzed. In our data there is a significant difference in the scale of relations in various sectors of the population under study, that is, between the central and the peripheral segment. As Romo (1986) and Romo and Anheier (1991) have pointed out, social networks are often composed of two general components: one component (the  $\alpha$  set, which is here the central segment) contains a number of small and relationally active (i.e., exhibits a large number of sent and received ties) blocks; the other component (the  $\Omega$  set, which is here the peripheral segment) often contains over half of the population clustered into one or two blocks that display a comparatively low incidence of sent and received ties. In a typical ICON-H analysis, the  $\alpha$  set produces numerous structurally equivalent blocks characterized by complex and meaningful relational interlock. The  $\Omega$  set, in contrast, produces no more than one or two very large blocks characterized by what appears as an almost random scattering of the few sent and received ties that they possess. Romo and Anheier's (1991) analysis of the data on German writers shows that members of the  $\Omega$  set, such as the peripheral segment reported here, are likely to be outside the information loop characterizing the social

organizational milieus within which they operate. Thus, they do not have adequate information about other members of the population to make very many interpersonal choices of others, nor do other members of the population have adequate information about members of the  $\Omega$  set to include them in their interpersonal choices. This results in a difference in the incidence scale of sent and received ties.

This difference in scale requires hierarchical blockmodeling. If the algorithm is allowed to meander through the data seeking an optimal split, it is likely to find one in the area of the network that contains the greatest incidence of ties; that is, somewhere among the  $\alpha$  set. As mentioned above, at some point, this produces single-member blocks (what we have termed *tp*). In order to obtain splits of the large, relationally inactive  $\Omega$  blocks, a hierarchical extension has been built into the ICON-H algorithm that holds constant all other blocks while it splits the target block.

#### *ICON-H and Other Blockmodeling Techniques*

In applying the ICON-H to data characterized by such a large-scale difference in the incidence of relations, it is typical that the  $\Omega$  set is clustered into a single block at a relatively low level of refinement. Solving the blockmodel for higher levels of refinement results in a greater number of blocks among members of the  $\alpha$  set while the  $\Omega$  block remains relatively unchanged. One solution to this problem is to use factor-analytical techniques (such as CONCOR) that induce clusters on data transformed into pairwise similarity measures based on *Z*-scores (product-moment correlations). For our purposes, however, product-moment correlations camouflage an important component of the structure we seek to investigate. Not only are we interested in the various aggregate patterns of sent and received ties, we are also interested in how such patterns interact with the incidence or frequency of sent and received ties to form structurally equivalent sets. Clearly, pairwise correlations, which are based on the products of *Z*-scores, substantially constrain the variability in the incidence of sent and received ties found in the original data. Given the goals of our analysis, this condition is unacceptable.

Nonetheless, another advantage of the CONCOR method is that some implementations are interactive and permit the researcher to select the block that will be subjected to further refinement (i.e., splitting). Thus, researchers can target larger blocks for additional refinement while holding other blocks constant. Other blockmodeling methods (e.g., Ward's method and single linkage, both implemented by Burt's popular STRUCTURE software) induce blocks by optimizing some object target function that presumably operationalizes the structural equivalence principle.

These methods produce blockmodels based on objective criteria, while blockmodels produced by CONCOR can be much more discretionary.

While we reject the use of product-moment correlations as similarity measures, we do like the discretion that an interactive implementation of CONCOR offers. Nonetheless, we are equally attracted to the idea of a baseline blockmodel purely based on optimizing an objective operationalization of the structural equivalence principle. Our strategy in this analysis is to do both. First we use the ICON-H objective target function (see eqq. A1–A4) to induce the six-block-level model. This model is composed of four elite and subelite blocks (the  $\alpha$  set), and the semiperiphery and periphery blocks (the  $\Omega$  set). Then we apply the interactive hierarchical capabilities of the ICON-H method to further partition the periphery.

Finally, Romo and Anheier (1994) have compared ICON-H to the other popular variance-minimizing clustering techniques on over 40 network data sets taken from the Project Metropolitan data archives (Janson 1975). The algorithms they studied include Ward's method, complete link, average link, and single link. Using a goodness-of-fit measure based on one suggested by Noma and Smith (1985), they found that the ICON-H method produced better fitting solutions than any of the other algorithms. Among the data sets used by Romo and Anheier (1994) was the data from the German writers' networks. In table A1 below, we reproduce this part of their analysis.

Noma and Smith (1985) use the squared ratio of the between sum of squares to the total sum of squares as their measure of goodness of fit. As will become clear below, squaring this ratio is superfluous, and, in the case of sparse binary matrices, the squared ratio can obscure major differences between solutions. In table A1 therefore, we use the ratio without squaring it. After Romo and Anheier (1994), we call this the "*R*-statistic," and it can be interpreted as the amount of variation in the network data explained by the block densities. The *R*-statistic varies between zero and one (0 = indicates no fit between the modeled densities and the observed data; 1 = a perfect fit between modeled densities and the observed data).

Table A1 presents the *R*-statistic for each network, the overall between sum of squares, the grand *R*-statistic, and the block sizes for the six-block solutions induced by ICON-H, Ward's method, complete linkage, average linkage, and single linkage algorithms. This table indicates that, for these data, ICON-H provides the best-fitting six-block solution. Concentrating on the grand *R*-statistic, ICON-H and Ward's method solutions produced *R*-statistics that were very close in magnitude (.209 and .171, respectively). However, in the case of sparse binary matrices, this apparent small difference is deceptive. One can get a sense of the difference in the way the two algorithms assign individuals to blocks by considering



TABLE A1  
GOODNESS OF FIT FOR FIVE BLOCKMODELING METHODS

	ICON-H Method	Ward's Method	Complete Linkage	Average Linkage	Single Linkage
<i>R</i> -statistic:					
Network 1 .....	.306	.271	.259	.255	.200
Network 2 .....	.286	.302	.266	.269	.252
Network 3 .....	.190	.229	.221	.230	.162
Network 4 .....	.503	.429	.326	.310	.297
Between sum of squares .....	470.259	435 102	227.952	211 855	181.678
Grand <i>R</i> -statistic .....	.209	.171	.092	.085	.073
Block size:					
Block 1 .....	53	53	131	133	134
Block 2 .....	33	50	3	2	1
Block 3 .....	22	17	2	1	1
Block 4 .....	20	12	1	1	1
Block 5 .....	6	6	1	1	1
Block 6 .....	5	1	1	1	1

NOTE.—This table is drawn from analyses presented in Romo and Anheider (1994).

the block sizes. Note that we use block designations 1, 2, 3, . . . , rather than A, B, C, . . . , to avoid confusion with the block assignments reported in the main body of the text. While block 1 (equivalent to the periphery block and the light culture block in the analysis reported above) in both cases contains 53 individuals (indeed, they were the same individuals), other blocks differ considerably and by as many as 17 individuals (in block 2, the semiperiphery). When a case-by-case comparison is made, there are even greater differences. Turning to the network-specific *R*-statistics, note that ICON-H performs better than Ward's method on networks 1 (AWARENESS) and 4 (DINNER INVITATION), while Ward's method performs better on networks 2 (FRIENDSHIP) and 3 (ASSISTANCE). It is interesting to note that networks 1 and 4 are the densest (i.e., contain the most network ties) while networks 2 and 3 are the sparsest. This analysis identifies an evident weakness in Ward's method solutions: they are overdetermined by sparse matrices. This problem results from transforming the observed binary networks into a single Euclidean distance matrix. Rather than weighting the contribution of a network to the overall solution in terms of the number of ties it contains (i.e., its overall density), the Euclidean distance transformations induced by Ward's method tend to give equal weight to all matrices, irrespective of density. The main problem with Ward's method or any other algorithm that uses Euclidean distances (including complete, average, and single linkage methods) is that, unlike ICON-H, they were not designed to analyze multiple networks simultaneously. In single network problems, our experiences suggest that ICON-H and Ward's method may perform equally well, although we have not formally tested this hypothesis.

Table A1 indicates that three linkage methods perform poorly when compared to ICON-H and Ward's method. All three algorithms produced chained solutions: that is, they create many single-member blocks and one huge cluster of residuals (as illustrated by the block sizes).

In summary, we find the solutions from the ICON-H algorithm to be superior to solutions from other available methods for a number of reasons. ICON-H creates partitions that are not only sensitive to the patterns of sent and received ties, but also to variations in the absolute number or density of network ties. The CONCOR method, because it transforms network data into product-moment correlations, is insensitive to variation in the number of sent and received ties. ICON-H is also appropriate for multiple network problems whereas most of the algorithms that use Euclidean distance transformations are not. This is particularly the case with Ward's method, which gives solutions that are overdetermined by sparse networks. Finally, the ICON-H method is not as susceptible to chaining, whereas complete, average, and single linkage methods are.

REFERENCES

- Anheier, Helmut K., and Jürgen Gerhards. 1991. "Literary Myth and Social Structure." *Social Forces* 69 (3): 811-30.
- Arabie, Philipp, Scott A. Boorman, and Paul Levitt. 1978. "Constructing Blockmodels: How and Why." *Journal of Mathematical Psychology* 17:21-63.
- Back, Kurt W., and Donna Polisar. 1983. "Salons und Kaffeehäuser." Pp. 276-86 in *Gruppensoziologie*, edited by Friedhelm Neidhardt. Opladen: Westdeutscher Verlag.
- Becker, Howard S. 1982. *Art Worlds*. Berkeley and Los Angeles: University of California Press.
- Benzecri, Jean-Paul. 1992. *Correspondence Analysis Handbook*. New York: Marcel Dekker.
- Blau, Judith. 1989. *The Shape of Culture: A Study of Contemporary Cultural Patterns in the United States*. Cambridge: Cambridge University Press.
- Boorman, Scott A., and Paul R. Levitt. 1983. "Blockmodelling Complex Statutes: Mapping Techniques Based on Combinatorial Optimization for Analyzing Economic Legislation and Its Stress Points over Time." *Economic Letters* 13:1-10.
- Boorman, Scott A., and Harrison White. 1976. "Social Structure from Multiple Networks II. Role Structures." *American Journal of Sociology* 81:1384-1446.
- Bourdieu, Pierre. 1979. *La Distinction: Critique sociale du jugement*. Paris: Minuit.
- . 1983. "Ökonomisches Kapital, kulturelles Kapital, soziales Kapital." Pp. 183-98 in *Zur Theorie sozialer Ungleichheiten*, edited by Rheinhard Kreckel. Göttingen: Verlag Otto Schwarz.
- . 1985a. "The Market of Symbolic Goods." *Poetics* 14:13-44.
- . 1985b. "The Genesis of the Concepts of 'Habitus' and 'Field.'" *Sociocriticism* 2 (2): 11-24.
- . 1986. "The Forms of Capital." Pp. 241-58 in *Handbook of Theory and Research for the Sociology of Education*, edited by John G. Richardson. New York: Greenwood.
- . 1989. "Social Space and Symbolic Power." *Sociological Theory* 7 (1): 14-25.
- . 1990. *The Logic of Practice*. Stanford, Calif.: Stanford University Press.
- . 1992. *Les Regles de l'art: Genese et structure du champ litteraire*. Paris: Editions du Seuil.
- Bourdieu, Pierre, and Loic J. D. Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.
- Breiger, Ronald L., Scott A. Boorman, and Philipp Arabie. 1975. "An Algorithm for Clustering Relational Data: With Applications to Network Analysis and Comparison with Multidimensional Scaling." *Journal of Mathematical Psychology* 12: 328-83.
- Burt, Ronald S. 1982. *Toward a Structural Theory of Action*. New York: Academic Press.
- . 1987. "STRUCTURE." Technical Report TR2. Center for the Social Sciences, Columbia University.
- DeNooy, Walter. 1992. "Social Networks and Classification in Literature." *Poetics* 20 (5-6): 507-37.
- DiMaggio, Paul. 1986. "Structural Analysis of Organizational Fields: A Blockmodel Approach." *Research in Organizational Behavior* 8:335-70.
- . 1987. "Classification in Art." *American Sociological Review* 52:440-55.
- DiMaggio, Paul, and Walter W. Powell. 1983. "Institutional Isomorphism." *American Sociological Review* 48 (2): 147-60.
- Gartman, David. 1991. "Culture as Class Symbolization or Mass Reification? A Critique of Bourdieu's *Distinction*." *American Journal of Sociology* 97 (2): 421-47.
- Gerhards, Jürgen. 1986. "Die Vergesellschaftung des Künstlers in der Moderne am Beispiel des Kaffeehauses." *Sociologica Internationalis* 1:73-93.

- Gerhards, Jürgen, and Helmut K. Anheier. 1989. "The Literary Field: An Analysis of Bourdieu's Sociology of Art." *International Sociology* 4:131-46.
- Greenacre, Michael. 1984. *The Theory and Application of Correspondence Analysis*. London: Academic Press.
- . 1990. *SimCA 2: User's Manual*. Pretoria: University of South Africa, Department of Statistics.
- Janson, Carl-Gunnar. 1975. "Project Metropolitan: A Longitudinal Study of a Stockholm Cohort." *Project Metropolitan Research Report*, no 2. University of Stockholm, Department of Sociology.
- Johnson, S. 1967. "Hierarchical Clustering Schemes." *Psychometrika* 32:241-54.
- Kadushin, Charles. 1974. *The American Intellectual Elite*. New York: Little, Brown.
- . 1976. "Networks and Circles in the Production of Culture." *American Behavioral Scientist* 19:769-84.
- Kelly D. 1980. "Results Using Blockmodel Target Functions." Mimeograph. Yale University, Department of Sociology.
- Knoke, David, and Ronald S. Burt. 1983. "Prominence." Pp. 195-222 in *Applied Network Analysis: A Methodological Introduction*, edited by Ronald S. Burt and Michael J. Minor. Newbury Park, Calif.: Sage.
- Lorraine, Francois, and Harrison C. White. 1971. "Structural Equivalence of Individuals in Social Networks." *Journal of Mathematical Sociology* 1 (1): 49-80.
- Mannheim, Karl. 1940. *Man and Society in an Age of Reconstruction*. New York: Harcourt, Brace & World.
- Müller, Hans-Peter. 1985. "Kultur, Geschmack und Distinktion: Grundzüge der Kulturosoziologie Pierre Bourdieus." In *Kultur und Gesellschaft*, edited by Friedrich Neidhardt, M. Rainer Lepsius, and Johannes Weiss. Opladen: Westdeutscher Verlag.
- Noma, Elliot, and D. Randall Smith. 1985. "Benchmark for the Blocking of Sociometric Data." *Psychological Bulletin* 97 (3): 583-91.
- Peterson, Richard A. 1985. "Six Constraints on the Production of Literary Works." *Poetics* 14:45-67.
- Romo, Frank. 1980. "Durkheim's Division of Labor Revisited: Topics in Structural Sociology." Mimeograph. Yale University, Department of Sociology.
- . 1986. "Moral Dynamics: A Blockmodeling Study of Conflict in a Mental Hospital." Ph.D. dissertation. Yale University, Department of Sociology.
- Romo, Frank, and Helmut K. Anheier. 1991. "The Omega Phenomenon: A Study of Social Choice." Working Paper. Russell Sage Foundation.
- . 1994. "Comparing Network Algorithms: Combinatorial and Distance-Based Optimizations." Working Paper Series of the Institute for Social Analysis. State University of New York, Stony Brook.
- Sailer, Lee D. 1978. "Structural Equivalence. Meaning and Definition, Computation and Application." *Social Networks* 1 (1). 73-90.
- Wasserman, Stanley, and Joseph Galaskiewicz. 1993. "Social Networks Analysis: Concepts, Methodology, and Directions for the 1990s." *Sociological Methods and Research* 22 (1): 3-22.
- Weber, Max. 1972. *Wirtschaft und Gesellschaft*. Tübingen: Mohr.
- Weller, Susan, and A. Kimball Romney. 1990. "Metric Scaling: Correspondence Analysis." *Sage University Paper Series on Quantitative Applications in the Social Sciences*, no. 07-075. Newbury Park, Calif.: Sage.
- White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structures from Multiple Networks: I. Models of Roles and Positions." *American Journal of Sociology* 81:730-80.

# Durkheim's Theories of Deviance and Suicide: A Feminist Reconsideration<sup>1</sup>

Jennifer M. Lehmann  
*University of Nebraska*

In *The Division of Labor in Society*, Durkheim conceptualizes deviance as an essentially asocial phenomenon, and he conceptualizes "woman" as an essentially asocial being. Both theories contradict Durkheim's characteristic social determinism, and both encounter, in *Suicide*, two further contradictions. First, *Suicide* demonstrates conclusively that relatively asocial individuals, women, are actually much less prone to deviance than relatively social individuals, men. Second, *Suicide* introduces the theory that deviance is an essentially social phenomenon that is produced by pathological social forces or "currents" rather than by "excessive individualization" and "insufficient socialization." Durkheim's second theory of deviance thus simultaneously rescues his theory of the social nature of men and his theory of the asocial nature of women.

## INTRODUCTION

Perhaps the most compelling claim made by liberals on behalf of modern industrial capitalism is that it bestows human rights universally on the abstract human being, the modern human "individual." Perhaps the most compelling critique of modern industrial capitalism and its liberal doctrine is that "universal" human rights, such as economic opportunity, political power, and cultural expression, are systematically and effectively denied to the majority of human beings under its purview. It follows that an investigation of the relationship between "the individual" and "society" in modernity accepts the terms of the liberal problematic but also bears the potential of critiquing liberal society on its own terms, according to its own criteria.

The question of whether, in theory and in practice, modern industrial capitalism actually realizes its own individualistic ideals has been posed repeatedly, diversely, and with widely varying results. The most persis-

<sup>1</sup> I would like to thank the *AJS* reviewers. Correspondence may be addressed to Jennifer M. Lehmann, Department of Sociology, 741 Oldfather Hall, University of Nebraska, Lincoln, Nebraska 68588-0324. E-mail: jlehmann@unlinfo.unl.edu

tent critical answers to this question have explored the limits of liberal individualism as demonstrated in the exclusion from the social contract of distinct individuals according to group identities such as class, race, and sex. Thus, the debate about modernity inevitably turns equally on the liberal categories of individual-level freedom and equality and the critical categories of group-level subordination and exploitation.

The critical perspective tends to evaluate modern society against an alternative and as yet unrealized modern society that has been reformed or revolutionized to produce greater *de facto* liberty and equality. It is also possible to evaluate modern society against traditional or feudal society, so that one may assess the extent to which modernity and liberalism have fulfilled their promise to eradicate ascriptive social systems of caste and conservative social theories of natural, group-level difference and inequality. Marxists argue that industrial capitalism is predicated on economic, political, and cultural inequality, notably on the basis of economic class. Multiculturalists argue that modern Westernism is Eurocentric and predicated on the internal and external colonization of non-European groups by Europeans. Feminists argue that liberal democracy coexists with feudal patriarchy and has been conjoined with sexual inequality, which is reflected in conservative social theory and patriarchal social practice, since its inception.

At stake in these theoretical and practical discussions is the issue of whether modern society and the modern world system can be procedurally reformed to be more inclusive or whether they must be structurally transformed to be more egalitarian. Can inequalities of race, class, and sex be eliminated through equal opportunity and individual mobility in the process of competition for stratified positions? Or must they be addressed through equality of outcome, and the leveling of political, economic, cultural, sexual, racial, and national hierarchies? For the feminist debate about modernity, this issue devolves in large part on the nature of the relationship between capitalism and patriarchy.

According to the liberal feminist view, capitalism and patriarchy are inimical to each other. Caste systems in which social positions are allocated ascriptively and on a group basis, are seen as anomalous and anachronistic, irrational, inefficient, inequitable, unprofitable, and thus ultimately incompatible with capitalism. The meritocracy of equal opportunity and individual mobility is seen as rational, efficient, and equitable, and ultimately the natural, necessary distribution mechanism of a market economy, polity, and culture. In the liberal view, modern capitalism must, can, and will be reformed to include both women and men as generic "individuals" in both the public and private spheres.

According to Marxist, socialist, radical, and multicultural feminism, modern capitalism and patriarchal castes are inextricably linked. (For

descriptions of these feminist paradigms and readings by their proponents, see, e.g., Jaggar and Rothenberg [1993].) The public, social, male sphere, where capitalism, individualism, and liberalism prevail, is not incompatible with, but contingent on, a private, natural, female sphere, where patriarchy, caste, and conservatism prevail. Patriarchy is not a feudalistic survival that modernity will ultimately displace, but an intrinsic part of the modern, dual, social, and world systems. Caste systems based on sex, like caste systems based on class, race, ethnicity, culture, and nationality, are rational sources of reserve, low-paid, and unpaid labor and thus are rational sources of hyperexploitation and superprofits as well as conflict between and within subordinate groups. Therefore, caste systems such as patriarchy are not only compatible with but essential to capitalism and cannot be reformed or eliminated within its structural confines.

Within the context of this theoretical and practical debate among feminists about the relationship between capitalism and patriarchy, my article constitutes a feminist rereading of the sociology of Émile Durkheim. Modernity appears in Durkheim's work as a complex configuration of reformed capitalism and patriarchy. In fact, Durkheim's social theory actually comprises two social theories: a theory of the public sphere, capitalism, and men, and a theory of the private sphere, patriarchy, and women. The objective here is to determine the relationship between these two spheres—capitalism and patriarchy—and these two groups—men and women—as they coexist within Durkheim's thought. I hope that this determination can shed light on the coexistence of capitalism and patriarchy, men and women, within contemporary thought and reality.

Durkheim's dominant theoretical system represents and constructs the relationship between "the individual" and "society" as they confront each other in modernity. This theory of generic "individuals" is complex and ambiguous, but in many instances a predominant theoretical position can be identified and relative theoretical consistency can be established (Lehmann 1993). Durkheim's subordinate theoretical system represents and constructs the relationship between women and men as they confront each other in modernity. In general, Durkheim's theory of women, like his theory of individuals, is characterized by internal consistency (see Lehmann 1990, 1991, 1994). However, in all instances, Durkheim's theory of individuals and his theory of women are completely contradictory.<sup>2</sup>

<sup>2</sup> Commentators generally recognize the contradictory relationship between Durkheim's theory of women and his theory of individuals in society, and most agree that Durkheim depicts women as less social, or asocial, relative to men (see also Besnard

The consistent contradiction between the nature of individuals and the nature of women in Durkheim's theory of modernity is emblematic of the persistent contradiction between capitalism and patriarchy, individual mobility and caste, liberalism and conservatism in the reality of modernity. These contradictions, the theoretical and the real, can be resolved in two contrasting ways. The contradiction within Durkheim's theory can be attributed to the anomalous and anachronistic presence of patriarchy and thus can be interpreted as an artifact of his theoretical and social context and as an intrusion of ambient patriarchal ideology into his progressive theory of modernity. In this view, Durkheim's theory of women is simply erroneous and can simply be corrected. Theoretical contradiction can be transformed into theoretical consistency, if Durkheim's women are subsumed into the Durkheimian category of the individual. This revisionist corrective to the theoretical contradiction of patriarchy is equivalent to the reformist corrective to the ontological contradiction of patriarchy. In liberal theory and practice, the anomalous reality of patriarchy will be eliminated through the assimilation of women into the capitalist world of men.

An alternative resolution of the contradiction between Durkheim's theory of individuals and Durkheim's theory of women is to interpret Durkheim's "women" as women and Durkheim's "individuals" as men and Durkheim's dualistic theory as an accurate description of a dualistic modernity. According to this view, the so-called contradiction in Durkheim's theory is not erroneous, and not even contradictory. Durkheim's theory of the relationship of difference and inequality between men and women and the relationship of compatibility and complementarity between the dual spheres of capitalism and patriarchy is consistent with the contradictory nature of so-called modern reality. The dual Durkheimian configurations of organicism, social determinism, and liberal individual mobility with respect to men and conflict, biological determinism, and conservative caste ascription with respect to women reflect and shape the modern world and our modern theories about that world.

This article will examine Durkheim's theories of suicide as a site where

---

1973; Gane 1983, 1993; Jay 1981; Johnson 1972; Jordanova 1980; Kandal 1988; Lehmann 1990, 1991, March 1982; Pope 1976; Roth 1990; Sydie 1987; Tiryakian 1981; Wityak and Wallace 1981; Yeatman 1987). In addition to examining Durkheim's theory of women within the context of his own writing, it is instructive to examine it within the context of his historical social and intellectual milieu. I have done so elsewhere (Lehmann 1994). There are many excellent historical sources on this subject, among them Bell and Offen (1983), Boxer and Quataert (1978, 1987), Landes (1988), Lukes (1985), Offen (1984, 1986, 1988), Roth (1990), Tilly (1981), and Waelti-Walters and Hause (1994).



Durkheim discusses women relatively extensively and where he encounters numerous, symptomatic, theoretical difficulties. In *The Division of Labor in Society* Durkheim ([1893] 1933) delineates a coherent, dualistic model of sexual differences and dual spheres. In *Suicide* (Durkheim [1897] 1951) empirical findings concerning female suicide patterns contradict theoretical positions in *The Division of Labor in Society*, and Durkheim is forced to develop a specific new theory, female fatalism, to reconcile social facts and social theory. In "Divorce by Mutual Consent," an article written nine years later, Durkheim ([1906] 1978) retracts his theory of female fatalism, reinterprets his data, and neutralizes the troublesome, feminist, political implications of *Suicide*.<sup>3</sup>

### THE DIVISION OF LABOR IN SOCIETY

*The Division of Labor in Society* (hereafter *The Division of Labor*) focuses primarily on the division of labor among individuals in the public sphere, the sphere Durkheim identifies as "society." Durkheim postulates that the division of labor initially destroys mechanical social solidarity through individuation but ultimately creates organic social solidarity through interdependence. According to his organicist theoretical scheme, the social division of labor is necessary, beneficial, functional, and moral.<sup>4</sup>

The social division of labor is necessary because the social organism, like all natural life, is subject to the inexorable process of evolution. Durkheim interprets the European transformation from agrarian feudalism to industrial capitalism as an adaptive metamorphosis from simplicity to complexity. This transformation, interpreted as a division of labor,

<sup>3</sup> Besides these three works, Durkheim discusses women in modern society in *L'Année Sociologique*, most notably in his review of Marianne Weber (Durkheim [1910] 1978c). In later works Durkheim ([1898] 1963, [1912] 1965) concentrates on primitive society and describes radical differences and inequalities between primitive men and women. Gane (1983, 1993) provides an excellent discussion of Durkheim's theoretical shift or self-contradiction.

<sup>4</sup> "The division of social labor" would be a more accurate translation of *De la division du travail social*. However, I employ the term "social division of labor" for purposes of comparison with the term "sexual division of labor." It is important to note that, although Durkheim portrays society as the public sphere of economic and political activity from which women are absent, French women participated extensively in society thus defined. Offen (quoted in Boxer and Quataert 1987, p. 183) reveals that, since the mid-19th century, "France had led all other nations with the highest percentage of women active in the labor force." According to Offen, the percentage of French women in the paid labor force grew from 24% in 1866 to 38% in 1911. In 1896, Jeanne-E. Schmahl (in Bell and Offen 1983, p. 101) estimated that 6,000,000 French women were engaged in waged labor. Bell and Offen (1983), Boxer and Quataert (1987), Landes (1988), and Tilly (1981) provide vivid documentation of women's participation in political activities in Third Republic France.

is then interpreted as natural conformity to a universal natural law. Durkheim equates modernity with the structural differentiation and functional specialization among human individuals, conceived of as cells in a complex, superior social organism.

The social division of labor is also necessary because it solves an intractable problem: the natural and inevitable pressure on scarce resources created by the natural and inevitable process of population growth. The social division of labor is an unconscious, species-level, adaptive, evolutionary response to an environmental stimulus; fortuitously, it is an optimum response and engenders many beneficial, though unintended, consequences.

One consequence of the social division of labor is that it radically alters the way in which functional abilities are distributed and transmitted to structural elements in the social organism. According to Durkheim, in primitive society, functional abilities are distributed at the societal level. They are identical among all individuals, and they are transmitted genetically and without differentiation within each society. In feudal society, abilities are distributed at the group level. They are identical among all individuals within a particular caste, and they are transmitted genetically and without differentiation within each caste. In modern society, general abilities may be distributed at the group level and transmitted genetically; specific abilities are distributed at the individual level and are transmitted socially.

The social division of labor is obviously beneficial because structural differentiation and functional specialization are optimally efficient means of production. However, the supreme and moral benefit of the social division of labor, according to Durkheim, is that it creates functional interdependence and therefore eventually creates organic solidarity among the differentiated, specialized individuals in modern, divided society. Organic solidarity means that differentiated, specialized individuals are integrated with each other and regulated by each other, because they must engage in cooperative productive interactive processes and because they must exchange the products of their labor. Furthermore, organic solidarity implies organicism, or collectivism. Differentiated and specialized individuals are integrated with and regulated by society because they stand in relation to the social organism as single cells stand in relation to a complex biological organism. In modern society, individuals are completely dependent on the collective social organism, and the interests of individuals are completely utterly aligned with the interests of the collective social organism.

How do women fit into this model of modern society? There are two possible answers to that question. The most obvious answer is that women are human "individuals" who are structurally differentiated and

functionally specialized, and who are, consequently, organically integrated and organically regulated, in relation to other individual social cells, to occupational and institutional social organs, and to the collective social organism. The problem with this obvious answer is that it is not the one given by Durkheim in *The Division of Labor*.

In *The Division of Labor*, Durkheim introduces the concept "the sexual division of labor" in a section entitled, "Cases Where the Function of the Division of Labor Is to Bring Forth Groups Which Would Not Exist Without It" (see esp. pp. 56–61; see also pp. 122–23, 247, 250, 264–65, 320). According to Durkheim, "the sexual division of labor is the source of conjugal solidarity" (1933, p. 56). In primitive society, there is no sexual division of labor; therefore, "marriage is in a completely rudimentary state" and conjugal solidarity is "very weak." "On the contrary," Durkheim says, "as we advance to modern times, we see marriage developing. . . . It is certain that at the same time sexual labor is more and more divided. . . . Today, among cultivated people, the woman leads a completely different existence from that of the man. One might say that the two great functions of the psychic life are thus dissociated, that one of the sexes takes care of the affective functions and the other of intellectual functions" (1933, pp. 59–60).

Structural differentiation develops concomitantly with functional specialization between the sexes. As primitive society evolves into modern society, women remain primitive. "One may . . . see in the female form the aboriginal image of what was the one and only type from which the masculine variety slowly detached itself" (1933, p. 57). Conversely, men evolve, and thus become structurally differentiated from women. "With the progress of civilization the brain of the two sexes differentiates itself more and more . . . due both to the considerable development of masculine crania and to a stationary or even regressive state of female crania" (p. 60).

Just as the division of labor among individuals creates modern society and organic social solidarity, the division of labor between men and women creates conjugal society and organic conjugal solidarity. The role of the division of labor in each case is "to render societies possible which, without it, would not exist. . . . Permit the sexual division of labor to recede below a certain level and conjugal society would eventually subsist in sexual relations preeminently ephemeral. If the sexes were not separated at all, an entire category of social life would be absent" (1933, p. 61).<sup>5</sup> Durkheim apparently felt that the sexual division of labor in his

<sup>5</sup> This statement can be interpreted as a warning against feminist policies, as can this discussion "By constitution woman is predisposed to lead a life different from man. Nevertheless, there are societies in which the occupations of the sexes are in fact the same" (1933, pp. 264–65).

own era was sufficiently developed to ensure the existence of organic conjugal solidarity, and therefore, of conjugal society. "Conjugal solidarity, for example, such as today exists among the most cultivated people, makes its action felt at each moment and in all the details of life" (p. 61).

In fact, Durkheim uses the sexual division of labor as an exemplar for the social division of labor. Challenging the notion that the progress of the social division of labor could be finite, limited by the "organico-psychic constitution" of men, Durkheim invokes the extreme functional specialization within biological organisms and the ultimate structural differentiation between women and men: "Moreover, in the very order of psychic and social functions, has not the division of labor, in its historical development, been carried to the last stage in the relations of men and women? Have not there been faculties completely lost by both?" (p. 401n.2)

In *The Division of Labor*, then, Durkheim actually articulates not one but two "divisions of labor in society." The first is, properly, a social division of labor, in the sense that it pertains to social labor—that is, labor in the public sphere—and to social individuals, that is, men. I argue that the second division of labor is the structural differentiation and functional specialization between these social individuals, men, taken collectively, that is, society—and a monolithic Other—that is, woman.

The sexual division of labor is radically different from the social division of labor because it divides and unites two finite groups, men and women, as opposed to uniting and dividing an infinity of (male) individuals. Apart from this crucial difference, the sexual division of labor is strikingly similar to the social division of labor. Like the social division of labor, the sexual division of labor is, for Durkheim, necessary, beneficial, functional, and moral.

Like the social division of labor, the sexual division of labor reflects the universal law of evolution. The two sexes, like all individuals, are destined to become structurally differentiated and functionally specialized. In primitive society, the two sexes, like all individuals, are structurally and functionally identical. As the social organism evolves and develops, the individual cells that constitute society and that are constituted by society (men) necessarily evolve and develop concomitantly. The Other or woman, who is outside society and its influences, is not altered by social evolution and therefore remains the same while becoming increasingly different from men.

There are two primary implications of woman's static, unevolved, primitive nature. First, woman's functional abilities are distributed at the group level rather than at the individual level. The only differentiation and specialization woman is involved in is external—the differentiation and specialization between woman and men. "Woman" comprises

women, who are structurally undifferentiated and functionally unspecialized in relation to each other. Second, woman's functional abilities are transmitted genetically rather than socially. Woman cannot become, and does not need to be, individualized, socialized, or modern. In relation to each other, women constitute a caste. In relation to men, women constitute a caste.

Like the social division of labor, the sexual division of labor is not only necessary but beneficial. Primarily, the sexual division of labor functions to unite the two differentiated and specialized sexes in exchange, interdependence, and organic sexual or conjugal solidarity.

Durkheim's organicism always privileges the prerequisites of the collective social organism, and the reproduction of its individual cells, the most essential condition of its existence, is ensured through the compulsory heterosexuality of a sexual division of labor.<sup>6</sup> Furthermore, the traditional family that results from the sexual division of labor, serves to unite not only women and men, but through them the individual biological organism and the collective social organism. The traditional family, structured by the sexual division of labor, represents the intersection and integration of nature (women, reproduction, the private sphere) and society (men, production, the public sphere).

### SUICIDE

While objectionable from a feminist perspective, the theory of the sexual division of labor articulated in *The Division of Labor* is consistent and coherent.<sup>7</sup> Women are asocial beings, consigned to the private, domestic, familial sphere, and thus situated outside society. Men are social beings, interacting with each other, who constitute and are constituted by the social organism and the collective consciousness that together constitute society. The two sexes are divided by their respective structures and functions and united by their interdependence in organic sexual solidarity within a second, parallel society: conjugal society.

*Suicide* was published in 1897, just four years after *The Division of Labor*. In many ways, *Suicide* provides testing and validation of Durk-

<sup>6</sup> I am grateful to Roslyn W. Bologh for pointing out in conversation the heterosexist implications of Durkheim's thought, and to Adrienne Rich (1980) for the term, "compulsory heterosexuality."

<sup>7</sup> Of course, Durkheim's explanation of the "nature" or structure of women is a naturalistic one, and thus contradicts his consistent "sociologism": his rules of sociological method and his organicist principle of social determinism. However, if only social facts must be explained by social facts, and only social individuals are socially determined, the rules of sociological method would not necessarily apply to natural facts, and natural women are naturally determined.

heim's theories about primitive, transitional, and modern societies, as delineated in the earlier work. *Suicide* is equally an application and demonstration of Durkheim's methodology, as explicated in *The Rules of Sociological Method*. Thus, there is a great deal of theoretical congruity between *Suicide* and Durkheim's other major works. However, *Suicide* both reveals and creates several major contradictions within Durkheim's theory of society and between his theory of society and his theory of sexual difference. In particular, it is the data on women's suicide rates that betray the contradictions within the Durkheimian theoretical structure.<sup>8</sup>

In the first place, *Suicide* reveals a crucial but neglected contradiction in Durkheim's social theory, the contradiction between two alternative theories of deviance. Durkheim's theory of deviance is consistent to a point. In transitional and modern society, deviance occurs because of excessive individualization, or, conversely, insufficient socialization. The separate beings that emerge through the division of labor initially tend to be unintegrated and unregulated. Egoistic and anomic individuals tend to be antisocial, deviating from collective identity, collective interests, and collective norms and limits for individual behavior. Thus, individualism is responsible for social problems as well as for profound individual unhappiness. One extreme form of anomic and egoistic deviance is the object of Durkheim's study, suicide.<sup>9</sup>

The question remains: What causes individualism and individualistic deviance? Specifically, is individualistic deviance caused by individuals or by society? Durkheim's simplest answer to the problem of deviance is that it is caused by *individuals*, particularly those who are insufficiently integrated and regulated by society. The question then becomes: What is the nature of the individual? The individual primarily appears in Durkheim's work as a physical, biological, natural entity, a mere organism or body.

Durkheim's alternative answer to the problem of deviance is that it is caused by a society that insufficiently integrates and regulates individu-

<sup>8</sup> Durkheim's theory of women is most extensively developed and most frequently discussed in relation to his theory of suicide. See Besnard (1973), Cashion (1970), Johnson (1972), Johnson (1979), Kandal (1988), Lehmann (1990), Pope (1976), Sharma (1978), Tiryakian (1981), Wityak and Wallace (1981). In contrast, Durkheim's ([1888] 1978; [1910] 1978; [1921] 1965; 1980) sociology of the family has been relatively neglected; see Bynder (1969), Lehmann (1991), Lukes (1985).

<sup>9</sup> Durkheim chooses to study suicide because of all the forms of modern, individualistic deviance, it is the one most amenable to psychological explanation and thus serves as the most powerful test case of the sociological method. For Durkheim, suicide is a typical example of deviance; it is not an exceptional example or something other than deviance.

als. The question then becomes: How can society create its own problems? Several answers to this question can be found in Durkheim's work. First, society can cause individualism and social problems by producing the absence of social forces, and thus producing the presence of individualism, in transitional and modern society. Second, society can cause individualism and social problems by producing negative social forces, the social "currents" of egoism and anomie in transitional and modern society.<sup>10</sup>

Durkheim's theory of transitional and modern deviance is that it is caused by excessive individualism and insufficient socialization. These formulations elide the question of whether deviance is caused by individuals acting independently of society or by society acting through individuals. Similarly, Durkheim's (1951, p. 258) explanation of egoism and anomie leaves the question unresolved: "Both spring from society's insufficient presence in individuals."<sup>11</sup> Is society's insufficient presence in individuals absolute? Is it the abandonment of individuals to their own biological imperatives or their own ideas and behaviors? Or is society's insufficient presence in individuals relative? Is it the failure to counterbalance negative social forces with positive social forces? If Durkheim begs the question, his work in *Suicide* forces the issue.<sup>12</sup>

In addition to revealing an underlying contradiction in Durkheim's

<sup>10</sup> Besnard (1973) similarly describes two disparate theories of suicide in Durkheim's work: "le juste milieu" (the golden mean) or "la courbe en U" (the U curve) and "l'équilibre entre des forces contraires" (the balance of contrary forces) or "l'action de courants suicidogènes" (the action of suicidogenic currents) (pp. 34–36).

<sup>11</sup> Besnard (1973, p. 34) aptly calls this Durkheim's "philosophie du juste milieu" (philosophy of the golden mean). According to Durkheim's logic, there is an ideal, optimum level of socialization and an ideal, optimum level of individualization. Mechanical society produces excessive socialization and insufficient individualization, expressed as altruism (excessive integration) and fatalism (excessive regulation). Transitional society produces excessive individualization and insufficient socialization, expressed as egoism (insufficient integration) and anomie (insufficient regulation). Durkheim predicts that a golden mean between individualization and socialization will be produced when newly differentiated and specialized individuals are reintegrated and reregulated under conditions of organic solidarity. I agree with Besnard that Durkheim's concepts egoism/integration/altruism are distinct from his concepts anomie/regulation/fatalism. For contrasting views, see e.g., Johnson (1965) and Pope (1976).

<sup>12</sup> The related question, of what constitutes a balance between individualism and socialization—a balance presumably struck in modern society through organic solidarity—also remains open. Is this to be a balance between two genuine entities, the individual and society, or is it to be a balance between two social forces, individualism and socialization? I contend (Lehmann 1993) that Durkheim considers even the modern individual an entirely social product, and that "individualism" in the positive sense refers exclusively to structural differentiation and functional specialization—not to the existence of actual individual ideas or behaviors.

social theory, *Suicide* creates some contradictions in his otherwise coherent sexual theory. In particular, *Suicide* calls into question two central aspects of Durkheim's theory of women as it appears in *The Division of Labor*: the asocial nature of women and the organic sexual or conjugal solidarity ideally connecting women to men. Unlike the contradiction in Durkheim's social theory, which remains unresolved, the contradictions in his sexual theory are forced into provisional resolution by the cold, hard statistical facts in *Suicide*.

There are two specific empirical findings in *Suicide* that expose or generate serious theoretical problems for Durkheim. Both of these findings concern female rates of suicide.<sup>13</sup> In the first place, Durkheim ascertains that women's suicide rates are much lower than men's. Viewed in isolation, this statistic is neither surprising nor problematic. Viewed in the context of Durkheim's theories of women and his theories of deviance, it is quite disruptive. If women are asocial, as Durkheim asserts, and if unsocialized individuals tend to be deviant, as one of his theories of deviance asserts, then women should have higher rates of suicide than men. Conversely, men who may be relatively unsocialized compared to other individual men but who are as a group relatively socialized compared to women, should have lower rates of suicide than women.

In his second chapter on egoistic suicide (chap. 3), Durkheim claims that it is precisely because women are asocial relative to men that they have lower rates of suicide relative to men. In this chapter, Durkheim identifies the family—the presence and number of children—as a form of social integration that reduces the tendency to egoism and egoistic suicide. He then explains that, due to their asocial nature, women have no need for social integration and thus no tendency to egoism or egoistic suicide. In this respect, Durkheim compares women to “children,” “the aged,” “an animal,” “lower societies,” and “primitive man” (1951, p. 215).<sup>14</sup> “If women kill themselves much less often than men, it is because they are much less involved than men in collective existence; thus they feel its influence—good or evil—less strongly. So it is with old persons and children, though for different reasons” (p. 299).

There is a second empirical finding in *Suicide* that disrupts the theoretical edifice erected in *The Division of Labor*. Women and men are both affected by marriage, as indicated by the differential suicide rates associ-

<sup>13</sup> Durkheim's discussion of women in *Suicide* appears primarily in book 2, chaps. 3 and 5, and book 3, chap. 3. See esp. pp. 171–202, 215–16, 259–76, 384–86. See also pp. 70–72, 99, 118–19, 166–67, 299–300, 341–42, 354–55, 376–78.

<sup>14</sup> Durkheim's conception of women as asocial leads rather inevitably to a comparison between women and primitives. See esp. Gane (1983, 1993), Lehmann (1991), and Wityak and Wallace (1981).



ated with marital status. However, the effects of marriage are diametrically opposite for women and men. Specifically, marriage attenuates male suicide rates while it exacerbates female suicide rates. Durkheim demonstrates, for example, that childless married men have substantially lower rates of suicide than unmarried men (644 to 975 per million, respectively), while childless married women have substantially higher rates of suicide than unmarried women (221 to 150 per million, respectively). "From this table and the preceding remarks, it appears that marriage has indeed a preservative effect of its own against suicide. But it is very limited and also benefits one sex only" (p. 198).

Durkheim's interpretation of this statistical evidence is that, while marriage is beneficial for men, it is detrimental for women. In France at least, "conjugal society is harmful to the woman and aggravates her tendency to suicide." Conversely, there are "special benefits bestowed by marriage on the male sex" (p. 189). This theme is repeated throughout chapter 3. "Conjugal society, so disadvantageous for women, must, even in the absence of children, be admitted to be advantageous for men" (p. 193). Clearly, then, this finding problematizes Durkheim's theory of organic solidarity between women and men precisely at the site of that alleged unity of interests, "conjugal society." However, while the family functions to integrate potentially egoistic individuals, marriage functions to regulate potentially anomic individuals. Therefore, Durkheim defers his explanation of the effects of marriage to his chapter on anomic suicide.

It is in this chapter that Durkheim's theories of deviance, suicide, and women confront each other most dramatically. Here, Durkheim begins by examining the respective effects of divorce on the suicide rates of men and women. Once again, these effects are not merely different, but diametrically opposed. Specifically, the rate of divorce is directly related to the suicide rate of married men and inversely related to the suicide rate of married women. Durkheim sums up these statistical relationships in the form of a "law" that he describes as "beyond dispute": "*From the standpoint of suicide, marriage is more favorable to the wife the more widely practiced divorce is; and vice versa*" (p. 269; emphasis in the original).

Durkheim argues that divorce has contrasting effects on men and women because marriage has contrasting effects on them. "And indeed, marriage may very possibly act in an opposite way on husband and wife. For though they have the same object as parents, as partners their interests are different and often hostile" (p. 269). Having established the conflict of interest between men and women, concerning both marriage and divorce, Durkheim ventures to explain it. It is crucial that Durkheim attributes the sexually specific effects of marriage and divorce to the

central difference between the sexes: the difference between social and asocial beings.

Men are quintessentially social. Even their sexuality is social rather than natural, mental rather than physical. As an "intellectual" phenomenon, male sexuality is not subject to biological regulation and requires social regulation. For men, marriage provides such regulation, satisfying their potentially insatiable sexual desire by defining and limiting its object. In the absence of such social regulation, men are subject to conjugal or sexual anomie: a "morbid desire for the infinite" in the realm of sexual relations.<sup>15</sup> The possibility of divorce effectively deregulates sexual desire, producing sexual anomie and, ultimately, suicide. Sexual anomie, therefore, accounts for the higher rates of suicide among men in regions where divorce is practiced, as well as higher rates among unmarried men.

Women, however, are quintessentially asocial. Women's sexuality is natural and physical rather than social and mental, and it is therefore regulated—limited and satisfied—biologically. "Woman's sexual needs . . . are more closely related to the needs of the organism, following rather than leading them, and consequently find in them an efficient restraint. Being a more instinctive creature than man, woman has only to follow her instincts to find calmness and peace. . . . Marriage is not in the same degree useful to her for limiting her desires, which are naturally limited" (Durkheim 1951, p. 272).

For women, marriage represents unnecessary regulation. More important, it constitutes excessive regulation for them. "By fixing the conjugal state permanently, it prevents all retreat, regardless of consequences. By limiting the horizon, it closes all egress and forbids even legitimate hope" (p. 272). The excessive sexual regulation that women experience is diametrically opposed to the insufficient sexual regulation that men experience, and in this opposition Durkheim finds the solution to the riddles posed by his suicide data. "Speaking generally, we now have the cause of that antagonism of the sexes which prevents marriage favoring them equally: their interests are contrary; one needs restraint and the other liberty" (p. 274). Durkheim labels the excessive sexual regulation peculiar to women "fatalism":

"The above considerations show that there is a type of suicide the opposite of anomic suicide, just as egoistic and altruistic suicides are opposites. It is the suicide deriving from excessive regulation, that of

<sup>15</sup> For Durkheim, anomie has two dimensions: it is a lack of rules for behavior, primarily in the newly specialized economic sphere, and a lack of limits on aspirations, primarily material aspirations, in the economic and sexual sphere.

persons with futures pitilessly blocked and passions violently choked by oppressive discipline. . . . To bring out the ineluctable and inflexible nature of a rule against which there is no appeal, and in contrast with the expression 'anomy' which has just been used, we might call it fatalistic suicide" (p. 276n.25).

Sexual fatalism accounts for the higher rates of suicide among women in regions where divorce is prohibited as well as higher rates among married women. Since fatalism is restricted to primitives and women, Durkheim finds it uninteresting and relegates it to a footnote. Ironically, he compares the fatalistic suicides of women to the suicides of slaves and others suffering from "excessive physical and moral despotism" (p. 276 n.25).

Durkheim's theory of sexual anomie and fatalism explains the contrasting effects marriage and divorce have on women and men. It can also be invoked to help explain the fact that women's rates of suicide are universally and substantially lower than men's. If women, as asocial beings, are not unregulated, but biologically regulated, then they are not susceptible to anomie or anomic suicide. This is parallel to the notion that women, as asocial beings, do not need social integration and are not susceptible to egoism or egoistic suicide. Women, while outside the purview of social integration and regulation, do not need social integration and regulation and thus should be not more but less deviant than men.<sup>16</sup>

Durkheim's solution to the problem of women in *Suicide* has important implications for his conceptualization of deviance (1960). Specifically, it externalizes the individual/society dichotomy described in "The Dualism of Human Nature" (Durkheim [1914] 1960). Durkheim imputes the asocial, animalistic, "individualistic" dimension of human nature to women. But he divests the asocial, animalistic being of its antisocial, deviant tendencies. Asocial beings are not subject to socialization or social determinism, but they are subject to instinctual control and biological determinism.

At the same time, Durkheim imputes the social, mental, and moral dimension of human nature to men. But he invests the social, mental, and moral individual with antisocial, deviant tendencies. Social beings are subject to socialization and social determinism; but they are also subject to insufficient socialization, either as a relative absence of social forces or as the predominance of negative social forces over positive social forces. *Suicide*, then, inclines toward support for Durkheim's second theory of deviance. Deviance is the product of society working through

<sup>16</sup> Elsewhere, however, Durkheim pointedly attributes women's lower crime rates not to their morality, their feminine socialization, or even their biological regulation, but to their "lack of opportunity" (1951, p. 342).

absolutely social individuals, that is, men. It is not the product of autonomous individuals, which do not exist, or of absolutely asocial beings, that is, women. Like all social facts, deviance has a social cause; like all social facts, deviance does not apply to women.

Durkheim's answer to the question of women in *Suicide* has decisive implications for his conception of deviance in modern society. It has devastating implications for his conception of organic solidarity in conjugal society. Durkheim's theoretical maneuvering, his theory of sexual anomie and sexual fatalism, leaves two problems unresolved. In the first place, what are the policy implications of the fact that marriage appears functional for men and dysfunctional for women, while divorce appears functional for women and dysfunctional for men? Durkheim's initial answer is simple: "Must one of the sexes necessarily be sacrificed, and is the solution only to choose the lesser of the two evils? Nothing else seems possible as long as the interests of husband and wife in marriage are so obviously opposed" (1951, p. 384).<sup>17</sup>

In fairness to Durkheim, he does suggest another, long-term, solution to the problem of the sexual differentiation of suicidal etiology. Durkheim attributes the "antagonism" of interests between men and women to the social nature of men, which predisposes them to sexual anomie, and the natural nature of women, which predisposes them to sexual fatalism. Of marriage, Durkheim says: "It cannot simultaneously be agreeable to two persons, one of whom is almost entirely the product of society, while the other has remained to a far greater extent the product of nature" (1951, p. 385).

Durkheim consequently advises that women play "a more active and important" part in society. However, this solution is one that also causes more problems. First, Durkheim retains his theory that the structural differentiation and functional specialization between women and men is not only irreversible but also progressive in nature. "To be sure, we have no reason to suppose that woman may ever be able to fulfill the same functions in society as man. . . . The female sex will not again become more similar to the male; on the contrary we may foresee that it will become more different" (1951, p. 385). It is hard to imagine that, as women become ever more different from men, women's sexuality could somehow become more similar to men's sexuality.

Second, Durkheim qualifies women's "more active and important"

<sup>17</sup> The immediately preceding text is, "The only way to reduce the number of suicides due to conjugal anomy is to make marriage more indissoluble. What makes the problem especially disturbing and lends it an almost dramatic interest is that the suicides of husbands cannot be diminished in this way without increasing those of wives" (1951, p. 384).

part in society to the point that it becomes meaningless. Woman's part in society must be different from man's; it must be "peculiarly her own": "Both sexes would thus approximate each other by their very differences. They would be socially equalized, but in different ways" (1951, p. 385). In addition, Durkheim cautions that even legal equal opportunity would not serve to integrate the structure of "woman" into the functions of men. Should legal barriers to women's social participation be dismantled, natural barriers would remain. "She could choose more freely, but as her choice would be determined by her aptitudes it would generally bear on the same sort of occupations. It would be perceptibly uniform, though not obligatory" (1951, p. 385n.16).<sup>18</sup>

In *Suicide*, Durkheim recognizes that the sexual conflict of interest in relation to marriage and divorce undermines his own support of marriage and opposition to divorce. He takes sides in the conflict, supporting the interests of men and marriage at the expense of women and divorce. He also, unconvincingly, suggests means to attenuate the conflict of interest between women and men—as it specifically pertains to the effects of marriage and divorce, and in the unforeseeable, in fact unlikely, future. With respect to the more general and more immediate question of the relationship between women and men, women's interests and men's interests, Durkheim appears to accept a sexual conflict of interest as inevitable. His own findings in *Suicide*, combined with his overarching theory that women are "the product of nature," while men are "the product of society," destroys his thesis of organic sexual solidarity. In *Suicide*, sexual difference between social men and asocial women overrides sexual interdependence between differentiated and specialized members of conjugal society.

#### "DIVORCE BY MUTUAL CONSENT"

*The Division of Labor* establishes that women and men are structurally differentiated and functionally specialized. It further purports that sexual

<sup>18</sup> This is one of the most equivocal passages in all of Durkheim's work. In addition to the quotes in the text, compare the following statements about sexual difference and conflict: "It is by no means certain that this opposition must necessarily be maintained. Of course, in one sense, it was originally less marked than now, but from this we cannot conclude that it must develop indefinitely. For the most primitive social states are often reproduced at the highest stages of evolution, but under different forms, forms almost the opposite of their original ones. . . . Only when the difference between husband and wife becomes less, will marriage no longer be thought, so to speak, necessarily to favor one to the detriment of the other. As for the champions today of equal rights for woman with those of man, they forget that the work of centuries cannot be instantly abolished; that juridical equality cannot be legitimate so long as psychological inequality is so flagrant." (1951, pp. 385–86)

difference and the sexual division of labor are natural and that they are evolutionary. As such they are both functional and adaptive, necessary and beneficial, immutable and progressive. Women are asocial and affective by nature, predisposed to physical and emotional functions in the private sphere. Men are social and intellectual by nature, predisposed to mental and moral functions in the public sphere. The sexual division of labor is the social expression of natural sexual difference. More important, it serves an essential sociobiological purpose: it unites women and men in organic solidarity within conjugal society.

*Suicide* maintains the structural difference between asocial women and social men, and the functional specialization of women in the family and men in society. However, instead of creating sexual unity and solidarity, sexual difference and the sexual division of labor are shown to divide women and men in a sexual conflict of interest at the heart of conjugal society. The fact that marriage and divorce have opposite effects on women and men creates a theoretical problem for Durkheim. He solves it with his novel theoretical explanation: the dual constructs sexual anomie and sexual fatalism. Nonetheless, the fact remains that marriage is somehow detrimental to women, and divorce is somehow beneficial to women. This fact creates a political problem for Durkheim. He supports indissoluble marriage and opposes legalized divorce, policies that he shows have negative—even lethal—consequences for women. More generally, Durkheim supports the sexual division of labor, which he shows creates a conflict of interest, rather than organic solidarity, between women and men.

The theoretical and political problems posed by *Suicide* apparently continued to haunt Durkheim. He returned to the subject of marriage, divorce, sexual difference, and suicide nine years later, in an article entitled "Divorce by Mutual Consent" (Durkheim 1978b).<sup>19</sup> In this article, Durkheim reverses his own statistical analysis and retracts his own theoretical explanation of women's suicide rates. Actually, all of the problematic implications in *Suicide* are caused by two specific empirical findings, and Durkheim's reinterpretation of the data eliminates both of them. According to *Suicide*, marriage is associated with an increase in female suicide, and divorce is associated with a decrease in female suicide.

These two correlations alone imply that marriage is detrimental to

<sup>19</sup> Besnard (1973) gives this crucial article the attention it merits and provides a brilliant analysis of the article and its impact on Durkheim's theory of suicide. However, Besnard—like Johnson (1972), Pope (1976), and Tiryakian (1981), e.g.—tends to interpret Durkheim's theory of women as prejudice or a product of the times—i.e., as an error, a mistaken or unconscious intrusion of ideology or even ad hoc reasoning into his more methodical theory of society. Writing later, Gane (1983) is also familiar with "Divorce by Mutual Consent" and tends to believe, as I do, that Durkheim's sexual theory is a deliberate theoretical strategy, a reasoned and essential ancillary to

women, divorce is beneficial to women, and that each affects women and men in opposite ways. These two correlations alone prompted Durkheim to develop his theory of sexual anomie and fatalism. And it is these two correlations, with all of their implications and consequences, that Durkheim denies in "Divorce by Mutual Consent." Here, Durkheim asserts that the apparent effects of marriage and divorce on female suicide rates are precisely *apparent* effects.

"Divorce by Mutual Consent" consists of three sections. Sections 1 and 3 describe the negative effects of divorce and militate against the legalization of divorce by mutual consent. Section 2 addresses the woman question: "It is true that the preceding facts apply solely to men" (1951, p. 245). In *Suicide*, Durkheim had emphasized that divorce is beneficial for women, as evidenced by the lower rates of suicide among married women where divorce is practiced. In "Divorce by Mutual Consent," Durkheim asks, not how beneficial divorce is for women, but how detrimental it is: "Divorce does not appear unfavorable to married women" (1978b, p. 245). In addition, he has a new explanation for the fact that the suicide rate of married women in Paris, where divorce is prevalent, is lower relative to unmarried women than in the provinces, where divorce is rare. It is that unmarried women kill themselves more frequently in Paris than in the provinces. Thus, "the advantage enjoyed by married Parisian women is purely apparent."<sup>20</sup>

As he had done in *Suicide*, Durkheim once again relates the effects of divorce on women to the effects of marriage on them. However, in "Divorce by Mutual Consent," he produces an entirely different "general law" concerning the nature of these effects. "It does not seem that the practice of divorce affects feminine suicide in an appreciable way. Moreover, this fact should hold no surprise; it is a specific case of a more general law, which can be formulated as follows: the state of marriage has only a weak effect on the moral constitution of women" (1978, pp. 246-47).

---

his social theory. These alternative interpretations are contrasted in the conclusion below. Besnard and Gane concur that Durkheim's description of women excludes them from ostensibly universal categories, such as "l'homme," "l'humanité," "humans." I am grateful to an anonymous reviewer for the following pertinent, fair-minded, and well-stated observations: "The question of whether divorce by mutual consent was to be allowed by French law had become a contemporary political issue by 1906 in a way it was not in 1897; and this, in turn, reflected progress in the Third Republic's avowed policies of anti-clericalism and laicization, in which Durkheim played a major role."

<sup>20</sup> Durkheim speculates that young unmarried women in Paris are exposed to "special dangers" and commit suicide more frequently there "as a result of the conditions in which they live, or of the inherent weakness of their moral temperament, or for both reasons at once" (1951, p. 246).

In *Suicide*, Durkheim had emphasized that marriage is detrimental for women, as evidenced by the higher rates of suicide among married women than among unmarried women. In "Divorce by Mutual Consent," Durkheim asks, not how detrimental marriage is for women, but how beneficial it is; he responds that it is not beneficial. "This lack of effectiveness of conjugal society is particularly evident as far as suicide is concerned. When there are no children, married women seem to kill themselves somewhat more than unmarried women of the same age. . . . Since, therefore, marriage, in a general way, has only a slightly beneficial effect on her, it is quite natural that divorce has no very pronounced harmful effect on her. She stands somewhat beyond the moral effect of marriage" (1978, p. 247).<sup>21</sup>

In "Divorce by Mutual Consent," then, marriage does not have a positive effect on women's suicide rates, it has no effect on women's suicide rates. By implication, marriage has no effect on women, thus no detrimental effect on women, and no effect on women that is opposite to the effect of marriage on men. Similarly, divorce does not have a negative effect on women's suicide rates, it has *no* effect on women's suicide rates. By implication, divorce has no effect on women, thus no beneficial effect on women, and no effect on women that is opposite to the effect of divorce on men. Durkheim's new explanation of these new statistical relationships is actually an old explanation, reaffirmed and revitalized. Women are unaffected by marriage and divorce because marriage and divorce are social phenomena, and women, as asocial beings, are immune to all effects of all social facts.

Durkheim's statistical maneuvering in "Divorce by Mutual Consent" is slight, but its theoretical and political consequences are considerable. In the first place, Durkheim's suicide data no longer have any feminist implications. Women's suicidal tendencies no longer substantiate a critique of marriage, or an advocacy of divorce, based on the interests of women. Durkheim's opposition to legalized divorce is still a choice based on the interests of men and their society, but it no longer represents a choice between "two evils." If women are not affected by marriage and divorce, they have no interests with respect to marriage and divorce. If women have no interests with respect to marriage and divorce, there is no conflict of interest between women and men with respect to marriage and divorce. Durkheim forges a unity of interest between women and men by dissolving women's interests altogether.

In the second place, the contradiction of *Suicide* contained in "Divorce

<sup>21</sup> Notice that Durkheim now inexplicably interprets the fact that married women commit suicide more frequently than unmarried women to mean that marriage has "only a slightly beneficial effect on her."



by Mutual Consent" negates the contradiction of *The Division of Labor* contained in *Suicide*. In *The Division of Labor*, men and women are joined in organic solidarity within conjugal society. In *Suicide*, men and women are put asunder in a conflict of interest over conjugal society. In "Divorce by Mutual Consent" men and women are reunited in organic solidarity and unity of interest within and over conjugal society. Durkheim's theory of the sexual division of labor, and organic sexual solidarity, is restored. And Durkheim's theory of the relationship between men and women—at once a relationship of differentiation, specialization, and interdependence that mirrors the relationship among individuals (men) in (male) society, and a relationship that connects the private society of men and women to the public society of men—is tacitly reconstructed.

There is a specific discontinuity in Durkheim's theory of women, between *The Division of Labor* and *Suicide*, that "Divorce by Mutual Consent" rectifies. At the same time, there is a general continuity in Durkheim's theory of women that characterizes all three works—in fact Durkheim's work in its entirety. The elements that remain constant are as follows:

1. *Sexual difference*.—Women and men are structurally differentiated. Women are essentially asocial; men are essentially social.
2. *Sexual division of labor*.—Women and men are functionally specialized. Women specialize in asocial functions of reproduction in the private sphere; men specialize in social functions of production in the public sphere.

Durkheim's theory of women, whether characterized by sexual solidarity or sexual conflict, is a theory of separate spheres. No matter how the relationship between men and women is conceived, it is a relationship of difference. No matter how the relationship between the public sphere of social society and the private sphere of conjugal "society" is conceived, society and conjugal society are two different worlds.

## CONCLUSION

Durkheim's theory of women contradicts his theory of individuals in every facet of his social ontology. The key to the absolute dichotomy between individuals and women is based on the absolute dichotomy between social beings and asocial beings. Durkheim's contribution to social theory is his insistence that human individuals are socially determined. Conversely, his "contribution" to sexual theory is his insistence that women are precisely not socially determined. This dichotomy is, of course, a hierarchy. Asocial women are different from, other than, individuals; asocial women are inferior to social individuals.

In *The Division of Labor*, women's lack of social determination appears as a group-level structure or essential nature, which is primarily emotional and biological. This sex-specific structure predisposes women to specialize in affective and reproductive functions in the private, familial sphere. Thus women are incapable of individual differentiation and individual specialization in the intellectual and practical functions of the public, social sphere. Since women do not participate in the social division of labor, they do not participate in organic social solidarity. However, they do participate in the sexual division of labor and, therefore, in organic sexual solidarity. Women are linked to men, and the private, natural sphere is linked to the public, social sphere—society—through a second, subordinate silent “society,” conjugal society.

In *Suicide*, women's lack of social determination appears as their relative immunity from social problems and deviance. They are asocial, biological beings, yet they do not suffer from the diseases of insufficient socialization and excessive individualism: egoism and anomie. They are asocial, biological beings, thus they do not suffer from the predominance of negative social forces, which lead to egoism and anomie, over positive social forces, which lead to integration and regulation. They are asocial, biological beings, thus they suffer from fatalism when social regulation penetrates the natural realm of conjugal society in the form of indissoluble monogamous marriage.

In “Divorce by Mutual Consent,” women's lack of socialization appears as their complete immunity from all social forces. Women are not affected less than, more than, or differently than men by social forces, as their relatively lower and opposite rates of suicide would suggest. Instead, women are totally unaffected by social forces, as their absolutely low and invariable rates of suicide would suggest. Women are “individuals” in the pejorative sense of the term, as biological bodies. They are not, however, subject to individualistic deviance. Women are not “individuals” in the positive sense of the term, as social personalities. Therefore, they are not subject to socially induced maladies. While women do not suffer or cause social problems in their proper place, the implication is that they would both suffer and cause social problems in the public sphere, as men presumably do, to the extent that they resemble women—individual organisms without social minds.

The relationship between Durkheim's women and Durkheim's individuals (I revert here to the sense of the term “individual” employed throughout the article, i.e., the positive, social sense) is a relationship of difference—a primary difference between the biological and the social, and secondary differences between emotional and intellectual abilities, reproductive and productive functions, private and public spheres, biological and social problems and solutions, sexual and social conflict and

solidarity. Durkheim's theory of women seems to contradict his theory of individuals. Alternatively, it seems to complement his theory of individuals.

The difference between contradiction and complementarity is the difference between two views of Durkheimian theory and two views of modern society. According to the first view, Durkheim's theory of women is an anomaly, a contradictory exception to his modernist liberalism. Similarly, the persistence of sexual castes in modern society is an anomaly, a contradictory exception to the structural tendency of capitalism toward individual mobility and equal opportunity. According to the second view, Durkheim's theory of women is both accurate and prescient. Durkheim ingeniously describes and advocates an actual dualistic system of sexual specialization and separate spheres. Durkheim correctly comprehends that capitalism is compatible with, perhaps even contingent upon, the persistence of sexual and other castes.

The first view epitomizes a liberal, revisionist approach to Durkheim's work and a liberal, reformist approach to modern society. From this perspective, Durkheim's theory of women and his theory of individuals could both be corrected. Women could disappear from the private sphere, disappear as women, disappear into the public sphere, and disappear into the abstract, universal, anonymous, and inclusive category, *individuals*. Similarly, modern capitalism could be corrected, to extend the privileges of individualism to more and more individuals, and to eliminate the vestiges of caste from more and more enclaves. Modern capitalism could be reformed, to be, in Durkheim's (1933, p. 435) words, "more like itself." Women and men could be assimilated in theory and integrated in reality.

In many important ways, the trajectory of capitalism is precisely the increasing inclusion of excluded groups into society: politically, culturally, and economically. For women in particular, this inclusion comprises political gains in the form of suffrage, equal opportunity legislation, and the legalization of divorce; cultural gains in the form of feminist influences on social theory and practice; and economic gains in the form of substantial paid labor force participation and the sexual integration of most occupations. These transformations in sociosexual structures were being wrought even as Durkheim wrote; to some extent, especially in *Suicide*, he acknowledges them.

Alternatively, the second view epitomizes a critical but preservationist approach to Durkheim's work, and a more radical, transformative approach to modern society. From this perspective, Durkheim's sexual theory of women and his social theory of individuals are complementary. His dual theories represent the dual realities of modern society: the organic interdependence of capitalism and patriarchy, public society and private

families, waged labor and unwaged work, masculine men and feminine women. More generally, the Durkheimian duality reflects the way in which capitalism contains castes—group-based inequalities of opportunity, structural inequalities of outcome, and the cultural forms that sustain these inequities—and individualism pertains to a minority of individuals.<sup>22</sup>

In many, equally important ways, patriarchy has survived capitalism, and women are still politically, culturally, and economically differentiated from and subordinate to men. Politically, although women vote, they do not hold political office or positions of power in proportion to their numbers. Culturally, while feminism has made inroads into social beliefs and behaviors, sexism and misogyny are pervasive in attitudes and actions alike. Economically, women have substantially lower rates of paid labor force participation than men. When they do work for pay, they earn substantially less money than men, they work a double day or second shift, and they are concentrated within the pink-collar regions and below the glass ceiling, in dual, split, and segmented labor markets, wherein occupations, firms, and industries are segregated according to sex.<sup>23</sup>

In summation, both readings of Durkheim's classical combination of liberalism and conservatism, and both readings of the late 19th-century combination of capitalism and patriarchy, can be instructive for contemporary theories of late 20th-century ideologies and realities. Both readings suggest the necessity of a complex reading of "modernity," and both readings suggest that the promises of modernity cannot be fulfilled without substantial social change.

<sup>22</sup> At the same time, Durkheim's theories exemplify the dual philosophy of modern society, the combination of liberal individualism with a conservative belief in castes, the way contemporary liberalism theorizes universal human rights and infinite individual differences, while making exceptions of (theorizing separately and differently, i.e., conservatively) entire groups on the basis of sex and sexuality, race and culture, class and nation. I view Durkheim's political position as a complex synthesis of conservatism, liberalism, and radicalism. The secondary literature is remarkably divided on this point; most recent work explores the progressive and radical implications of his thought; see, e.g., Fenton (1980, 1984), Gane (1983), and Pearce (1989).

<sup>23</sup> In 1992, women's civilian labor force participation was, at 57%, about 74% that of men's (76.6%). In 1991, women in the paid labor force continued to earn 70% of what men earned, a level not much higher than the 1920 level (63%; see U.S. Bureau of the Census 1993; Marini 1989). For a theoretical discussion of economic sexual inequality and labor market structure, see, e.g., O'Donnell (1984) and Rebitzer (1993). For a theoretical discussion and empirical evidence of women's income inequality and the sexual segregation of labor markets, see, e.g., Tomaskovic-Devey (1993) and Marini (1989). For a discussion of the second shift and the double day, see, e.g., Shelton (1992), Hartmann (1984a). For a discussion of the relationship between capitalism and patriarchy, see, e.g., Hartmann (1984a, 1984b) and Pateman (1988).

# REFERENCES

- Bell, Susan Groag, and Karen M. Offen, eds. 1983. *Women, the Family, and Freedom: The Debate in Documents*, vols. 1, 2. Stanford, Calif.: Stanford University Press.
- Besnard, Philippe. 1973. "Durkheim et les femmes ou le Suicide inachevé." *Revue française de Sociologie* 14:27-61.
- Boxer, Marilyn J., and Jean H. Quataert, eds. 1978. *Socialist Women: European Socialist Feminism in the Nineteenth and Early Twentieth Centuries*. New York: Elsevier.
- . 1987. *Connecting Spheres: Women in the Western World, 1500 to the Present*. New York: Oxford University Press.
- Bynder, Herbert. 1969. "Émile Durkheim and the Sociology of the Family." *Journal of Marriage and the Family* 31:527-33.
- Cashion, Barbara G. 1970. "Durkheim's Concept of Anomie and Its Relationship to Divorce." *Sociology and Social Research* 55:72-81.
- Durkheim, Émile. (1888) 1978a. "Introduction to the Sociology of the Family." Pp. 205-28 in *Émile Durkheim on Institutional Analysis*, edited by Mark Traugott. Chicago: University of Chicago Press.
- . (1893) 1933. *The Division of Labor in Society*. New York: Free Press.
- . (1897) 1951. *Suicide*. New York: Free Press.
- . (1898) 1963. *Incest: The Nature and Origin of the Taboo*. New York: Lyle Stuart.
- . (1906) 1978b. "Divorce by Mutual Consent." Pp. 240-52 in *Émile Durkheim on Institutional Analysis*, edited by Mark Traugott. Chicago: University of Chicago Press.
- . (1910) 1978c. "Review of Marianne Weber, *Ehefrau und Mutter in der Rechtsentwicklung*." Pp. 139-44 in *Émile Durkheim on Institutional Analysis*, edited by Mark Traugott. Chicago: University of Chicago Press.
- . (1912) 1965a. *The Elementary Forms of the Religious Life: A Study in Religious Sociology*. New York: Free Press.
- . (1914) 1960. "The Dualism of Human Nature and Its Social Conditions." Pp. 325-40 in *Émile Durkheim, 1858-1917*, edited by Kurt H. Wolff. Columbus: Ohio State University Press.
- . (1921) 1965b. "The Conjugal Family." Pp. 528-36 in "A Durkheim Fragment," with an introduction by George Simpson. *American Journal of Sociology* 70:527-36.
- . 1980. *Contributions to L'Année Sociologique*, edited by Yash Nandan. New York: Free Press.
- Fenton, C. Stephen. 1980. "Race, Class and Politics in the Work of Émile Durkheim." Pp. 143-82 in *Sociological Theories: Race and Colonialism*. Paris: Unesco.
- . 1984. "Race and Society: Primitive and Modern." Pp. 116-42 in *Durkheim and Modern Sociology*, edited by Steve Fenton. Cambridge: Cambridge University Press.
- Gane, Mike. 1983. "Durkheim: Woman as Outsider." *Economy and Society* 12: 227-70.
- . 1993. *Harmless Lovers? Gender, Theory and Personal Relationships*. New York: Routledge.
- Hartmann, Heidi. 1984a. "The Family as the Locus of Gender, Class, and Political Struggle: The Example of Housework." Pp. 341-57 in *Feminist Frameworks*, edited by Alison M. Jaggar and Paula S. Rothenberg. New York: McGraw Hill.
- . 1984b. "The Unhappy Marriage of Marxism and Feminism: Towards a More Progressive Union." Pp. 172-89 in *Feminist Frameworks*, 2d ed. Edited by Alison M. Jaggar and Paula S. Rothenberg. New York: McGraw Hill.
- Jay, Nancy. 1981. "Gender and Dichotomy." *Feminist Studies* 7:38-56.

- Johnson, Barclay. 1965. "Durkheim's One Cause of Suicide." *American Sociological Review* 30:875-86.
- . 1972. "Durkheim on Women." Pp. 164-67 in *Woman in a Man-Made World*, edited by Nona Glazer-Malbin and Helen Youngelson Waehrer. Chicago: Rand McNally.
- Johnson, Kathryn K. 1979. "Durkheim Revisited: 'Why Do Women Kill Themselves?'" *Suicide and Life-Threatening Behavior* 9:145-53.
- Jordanova, L. J. 1980. "Natural Facts: A Historical Perspective on Science and Sexuality." Pp. 42-69 in *Nature, Culture and Gender*, edited by Carol P. MacCormack and Marilyn Strathern. New York: Cambridge University Press.
- Kandal, Terry R. 1988. *The Woman Question in Classical Sociological Theory*. Miami: Florida International University Press.
- Landes, Joan B. 1988. *Women and the Public Sphere in the Age of the French Revolution*. Ithaca, N.Y.: Cornell University Press.
- Lehmann, Jennifer M. 1990. "Durkheim's Response to Feminism: Prescriptions for Women." *Sociological Theory* 8 (2): 163-87.
- . 1991. "Durkheim's Women: The Durkheimian Theory of the Structures and Functions of Sexuality." *Current Perspectives in Social Theory* 11:141-67.
- . 1993. *Deconstructing Durkheim: A Post-post-structuralist Critique*. New York: Routledge.
- . 1994. *Durkheim and Women*. Lincoln: University of Nebraska Press.
- Lukes, Steven. 1985. *Émile Durkheim: His Life and Work*. Stanford, Calif.: Stanford University Press.
- March, Artemis. 1982. "Female Invisibility in Androcentric Sociological Theory." *Insurgent Sociologist* 11 (2): 99-107.
- Marini, Margaret Mooney. 1989. "Sex Differences in Earnings in the United States." *Annual Review of Sociology* 15:343-80.
- O'Donnell, Carol. 1984. "Major Theories of the Labour Market and Women's Place Within It." *Journal of Industrial Relations* 26 (2): 147-65.
- Offen, Karen M. 1984. "Depopulation, Nationalism, and Feminism in Fin-de-siècle France." *American Historical Review* 89 (3): 648-76.
- . 1986. "Ernest Legouvé and the Doctrine of 'Equality in Difference' for Women: A Case Study of Male Feminism in Nineteenth-Century French Thought." *Journal of Modern History* 58:452-84.
- . 1988. "Defining Feminism: A Comparative Historical Approach." *Signs* 14 (1): 119-57.
- Pateman, Carole. 1988. *The Sexual Contract*. Stanford, Calif.: Stanford University Press.
- Pearce, Frank. 1989. *The Radical Durkheim*. Boston. Unwin Hyman.
- Pope, Whitney. 1976. *Durkheim's Suicide: A Classic Analyzed*. Chicago: University of Chicago Press.
- Rebitzer, James B. 1993. "Radical Political Economy and the Economics of Labor Markets." *Journal of Economic Literature* 31:1394-1434.
- Rich, Adrienne. 1980. "Compulsory Heterosexuality and Lesbian Existence." *Signs* 5 (4): 647-50.
- Roth, Guenther. 1990. "Durkheim and the Principles of 1789." *Telos* 82:71-88.
- Sharma, Arvind. 1978. "Émile Durkheim on Suttie as Suicide." *International Journal of Contemporary Sociology* 15:283-91.
- Shelton, Beth Anne. 1992. *Women, Men and Time: Gender Differences in Paid Work, Housework and Leisure*. New York: Greenwood Press.
- Sydie, R. A. 1987. *Natural Women, Cultured Men: A Feminist Perspective on Sociological Theory*. New York: New York University Press.
- Tilly, Louise A. 1981. "Women's Collective Action and Feminism in France, 1870-1914." Pp. 207-31 in *Class Conflict and Collective Action*, edited by Louise A. Tilly and Charles Tilly. Beverly Hills, Calif.: Sage.

- Tiryakian, Edward A. 1981. "Sexual Anomie, Social Structure, Societal Change." *Social Forces* 59:1025-53.
- Tomaskovic-Devey, Donald. 1993. "The Gender and Race Composition of Jobs and the Male/Female, White/Black Pay Gaps." *Social Forces* 72 (1). 45-76.
- U.S. Bureau of the Census. 1993. *Statistical Abstract of the United States: 1993*. Washington, D.C.: U.S. Bureau of the Census.
- Waelti-Walters, Jennifer, and Steven C. Hause. 1994. *Feminisms of the Belle Epoque: An Historical and Literary Anthology*. Lincoln: University of Nebraska Press.
- Wityak, Nancy L., and Ruth A. Wallace. 1981. "Durkheim's Non-Social Facts about Primitives and Women." *Sociological Inquiry* 51.61-7.
- Yeatman, Anna. 1987. "Women, Domestic Life and Sociology." Pp. 157-72 in *Feminist Challenges: Social and Political Theory*, edited by Carole Pateman and Elizabeth Gross. Boston. Northeastern University Press.

# Models for Describing the Underlying Structure of Sex Segregation<sup>1</sup>

Maria Charles

*University of California, San Diego*

David B. Grusky

*Stanford University*

This article introduces a structural approach to analyzing sex segregation data that rests on margin-free measures of the underlying association in sex-by-occupation arrays. The starting point for the analyses is a log-multiplicative model that is formally consistent with the conventional practice of summarizing cross-national variability in a single parameter pertaining to the overall strength of sex segregation. Under this baseline specification, the segregation regime is forced to take on the same basic shape in each country, with the only form of permissible variability being a uniform compression or expansion of the peaks and valleys characterizing the shared segregation profile. Although the latter model does not account for the cross-national variability in our illustrative data, it can be readily generalized in ways that both improve the fit and yield new insights into the structure and sources of sex segregation. These elaborated models can be used to examine the hierarchical structure of segregation, to identify the dominant "segregation profiles" in industrial countries, and to parse out the net residue of segregation at multiple levels of analysis.

The study of occupational sex segregation appears to be entering its take-off period. This can be seen, for example, in the recent resurgence of interest in describing how the structure of sex segregation varies across

<sup>1</sup> Earlier versions of this paper were presented to the American Sociological Association, Research Committee 28 of the International Sociological Association, the Working Group on Problems of the Low-Income Population, the Department of Sociology at the University of California—Los Angeles, and the Stanford Center for the Study of Families, Children, and Youth. We acknowledge research support from the National Science Foundation Presidential Young Investigator Program (NSF SES-8858467), the Stanford Center for the Study of Families, Children, and Youth (1992–93 grant), and the Stanford University Dean's Research Fund. This paper was written in part while Maria Charles was at the Swiss Federal Institute of Technology at Zurich and David B. Grusky was at the Center for Advanced Study in the Behavioral Sciences on fellowship leave funded by the NSF (BNS-8700864). We are grateful for the com-



nations, over time, and among industries, organizations, or economic sectors (e.g., Reskin 1993; Brinton and Ngo 1991, 1993; Hakim 1992; Presser and Kishor 1991; Baron, Mittman, and Newman 1991; Jacobs 1989a, 1989b; Tienda and Ortiz 1987; Tienda, Smith, and Ortiz 1987; Bianchi and Rytina 1986; Bielby and Baron 1984). What makes this takeoff period so distinctive is that it is unfolding *without* any major methodological innovations of the kind that have historically played transformative roles in other subfields of stratification research. The methods deployed by sex segregation researchers have, in fact, remained largely unchanged over the last 30 years, with the index of dissimilarity ( $D$ ) and its not-so-distant cousins still playing a featured role. At regular intervals, the relative merits of competing indices are ritually debated (see, e.g., James and Taeuber 1985; Massey and Denton 1988; Coulter 1989; Watts 1992), yet such commentary appears to have little influence on the subsequent conduct of sex segregation research. We see no evidence, moreover, of an emerging interest in modeling segregation data; indeed, whereas stratification researchers in other subfields have long since abandoned  $D$  (and other indices) in favor of model-based "structural parameters" (see Duncan 1984, p. ix; Sobel, Hout, and Duncan 1986), this standard methodological transition has not yet occurred among sex segregation researchers.

It is not unusual for research subfields to bear the imprint of the time period in which they initially became popular. In the case of segregation research, the first burst of activity occurred precisely when  $D$  was emerging victorious from a "ten-year index war" (see Peach 1975, p. 3), and we might therefore expect this index to have shaped the development of segregation theorizing and research in decisive ways. As it turns out, the index of dissimilarity and its close analogues quickly became entrenched in the field, so much so that the rise of log-linear modeling went largely unnoticed despite its (seemingly obvious) relevance to tabular analyses of sex segregation. The purpose of the present article, then, is to document some of the advantages of log-linear and log-multiplicative models in describing, comparing, and explaining patterns of occupational sex segregation. After introducing a general multiplicative framework for modeling sex-by-occupation tables, we will derive a new scalar index of sex segregation and specify the conditions under which such an index can satisfactorily represent cross-national variability. We will also introduce simple multiplicative models that can account for the sex-by-occupation

---

ments of Karen Aschaffenburg, James Baron, Marlis Buchmann, Diane Burton, Lisa Catanzarite, Clifford Clogg, Thomas DiPrete, Leo Goodman, Joon Han, Robert Hauser, Jerald Herting, Michael Hout, Robert Mare, Manuela Romero, Jesper Sørensen, Kimberly Weeden, and Yu Xie. Direct correspondence to Maria Charles, Department of Sociology, University of California, San Diego, California 92093-0102.

association in terms of exogenous variables defined at the level of occupations.

#### MARGINAL DEPENDENCIES IN SEGREGATION INDICES

It may be instructive to briefly review the methodological state of affairs among sex segregation researchers. As we have already noted, the starting point for most analyses is the index of dissimilarity, where this is defined as

$$D = \sum_{j=1}^J |(F_j/F) - (M_j/M)| \times 100 \times 1/2. \quad (1)$$

We have followed convention (see Duncan and Duncan 1955) in using  $J$  to refer to the total number of occupations,  $M_j$  and  $F_j$  to refer to the number of men and women in the  $j$ th occupation, and  $M$  and  $F$  to refer to the number of men and women in the labor force as a whole. The value of  $D$  can be interpreted, therefore, as the percentage of the labor force that must change occupations to bring about a perfect correspondence between the sex ratio within each occupation and the overall rate of female labor force participation (see Winship [1977, p. 1061] for further details and qualifications).

The flaws of  $D$  are certainly well known (e.g., James and Taeuber 1985), yet it seems that the full implications of the known have not been sufficiently appreciated. What must of course be stressed is that  $D$  is not invariant under multiplicative transformations of the occupational margins; as a result, it becomes difficult to interpret cross-national or temporal variability in  $D$ , since the driving force behind such variability can be simple distributional differences in the occupational structure as well as real heterogeneity in the sex composition of occupations (cf. Butler 1987; Silber 1989). It is here, then, that the heritage of  $D$  as an index of *residential* segregation reveals itself in an unfortunate way. Although most scholars treat occupational and residential segregation as direct analogues, it is important to recognize that the former is measured with categories that are standardized across the units of comparison (i.e., occupations), whereas the latter is typically calculated in terms of "tracts" or "districts" that are defined in highly idiosyncratic ways.<sup>2</sup> It is clearly

<sup>2</sup> In carrying out temporal comparisons, the logic underlying the analysis of occupational and residential segregation is much the same, since in both cases the categories of the classification system (i.e., occupations or census tracts) are largely unchanged over time, thereby making it possible to map each of the categories within the benchmark period into one, and only one, of the categories appearing subsequently. The same type of one-to-one mapping is also feasible when one compares occupations across spatial units (e.g., cities, nation-states), but not when one compares census tracts or districts across such units.

## OCCUPATIONAL COMPOSITION

		Margin Dependent	Margin Free
L	Margin Dependent	Isolation Index; Interaction Index; Variance Ratio Index	Size-Standardized Dissimilarity Index
A			
B			
O			
R	Margin Free	Atkinson Index; Gini Index; Index of Dissimilarity; Lieberson's Diversity Index	Odds Ratios (or functions of them)
F			
O			
R			
C			
E			

FIG. 1.—Assorted segregation measures cross-classified by two forms of marginal dependence. The interaction and isolation indices are often denoted  $xP^*y$  and  $xP^*x$  (see Lieberson 1981). The variance ratio index has been labeled  $S$  (Zoloth 1976),  $\eta^2$  (Duncan and Duncan 1955), and the “revised index of isolation” (Bell 1954). The diversity index cited here was introduced by Lieberson (1969), while the Atkinson index was recently reviewed by James and Taeuber (1985). See Coulter (1989) for a comprehensive review of related indices.

impossible to specify a one-to-one correspondence between the census tracts of different cities, and consequently the incentive for devising margin-free measures is somewhat reduced. In this regard, it would be unreasonable to expect the conventional measures of intergroup inequality to be margin free, since most of them were devised with the empirical case of residential segregation in mind (see fig. 1).

We do not mean to imply that contemporary segregation research is invariably carried out with  $D$  alone. In fact, there is a small industry of research based on the premise that marginal effects *should* be purged from the data, with the point of departure typically being some type of modified or corrected version of  $D$ . We are referring, for example, to the well-known proposal of Blau and Hendricks (1979) to decompose changes in  $D$  into components attributable to occupational restructuring and residual “shifts in sex composition” (p. 199; see also Fuchs 1975; England 1981; Handl 1984; Tienda and Ortiz 1987; Beller 1984; Bianchi and Rytina 1986). In more recent work, Abrahamson and Sigelman (1987) sought to purge  $D$  of marginal dependence by regressing it on the “structural propensity toward occupational segregation” (p. 591), while Bridges (1982) proposed to adjust  $D$  “based on a comparison of the observed

level of occupational segregation with that expected given the occupational mix" (p. 278). We would interpret work of this kind as expressing an incipient interest in margin-free measures of association; however, given that  $D$  has become so entrenched in the field, this interest was inevitably channeled into purely reformist efforts (i.e., modifying  $D$ ).

The latter approaches have attracted some attention (see, e.g., Stafford and Fossett 1989), but it would appear that the size-standardized index of dissimilarity ( $D_s$ ) is gradually becoming the de facto standard for comparative analyses of sex segregation (see Gibbs 1965; Gross 1968). We can calculate  $D_s$  in the following way:

$$D_s = \sum_{j=1}^J \left| \left[ (F_j/T_j) / \sum_{j=1}^J (F_j/T_j) \right] - \left[ (M_j/T_j) / \sum_{j=1}^J (M_j/T_j) \right] \right| \times 100 \times 1/2, \quad (2)$$

where  $T_j$  refers to the total number of males and females in the  $j$ th occupation (i.e.,  $T_j = M_j + F_j$ ). As indicated in equation (2),  $D_s$  will be unaffected by the shape of the occupational distribution, since it standardizes each of the  $J$  occupations to the same size. The two numerators in this equation (i.e.,  $[F_j/T_j]$  and  $[M_j/T_j]$ ) index the female and male proportions in the  $j$ th occupation, while the corresponding denominators calibrate these values against the proportions prevailing in other occupations. As the amount of sex segregation increases, the difference between the "calibrated proportions" grows large, and the value of  $D_s$  increases in tandem. This type of standardization has been recently applied by Brinton and Ngo (1991, 1993), Presser and Kishor (1991), Williams (1979), Jacobs (1989a, 1989b), and Jacobs and Lim (1992).

It turns out that standardizing  $D$  is far from cost free. While the usual standardization does eliminate one form of marginal dependence, it has the perverse effect of introducing a new dependence on the rate of female labor force participation.<sup>3</sup> The appeal of  $D$  has long been its "scale invariance" (James and Taeuber 1985, pp. 15–17); that is, the value of  $D$  is unaffected by simple multiplicative transformations of the sex ratio, and consequently it can safely be used to compare countries, cities, or time periods with differing rates of female labor force participation. However, the same property does not hold for  $D_s$ , with the implication being that this type of standardization is seriously flawed for purposes of comparative research. We might well regard  $D_s$  as the analogue to the "Rogoff index" of mobility research (Rogoff 1953; see also Hauser 1978), since it suffers from the same types of marginal dependence that doomed

<sup>3</sup> This can be readily demonstrated by example. If the number of females in each occupation is multiplied by an arbitrary constant, the value of  $D_s$  will typically change.

this index and ultimately ushered in a new era of log-linear modeling. The flaws of  $D$ , have *not* been sufficiently appreciated in the literature to date; if anything, it appears that  $D$ , is becoming increasingly popular, with several recent studies featuring it as one of the key indices of sex segregation (e.g., Jacobs 1989a).

The foregoing comments suggest that the current state of affairs is less than ideal. As indicated in figure 1, the index of dissimilarity is invariant under multiplicative transformations of the sex ratio but not under multiplicative transformations of the occupational margins. At the same time, the size-standardized index successfully eliminates the latter dependence, but only at the cost of losing the scale invariance that characterized the original index. If we wish to eliminate both forms of marginal dependence simultaneously, we have no alternative but to use measures that are functions of cross-product ratios (see Goodman 1991; Becker and Clogg 1989; Altham 1970). Although a few enterprising scholars have already applied log-linear models to sex segregation tables, none of the efforts to date has exploited the full potential of this approach (see Handl 1984; Willms 1982; Semyonov 1980; Semyonov and Scott 1983; Stolzenberg and D'Amico 1977). We think that further analyses based on margin-free measures of sex segregation may lead to nontrivial revisions of our understanding of cross-national and over-time variability in gender stratification.<sup>4</sup>

#### SCALAR MEASURES OF SEX SEGREGATION

We ought not forget that segregation indices are merely scalar summaries of complex "segregation curves" representing the sex composition of all occupations. To be sure, most segregation scholars are quick to concede that "no single measure is correct for all purposes" (e.g., Lieberman 1980, p. 253; Abrahamson and Sigelman 1987), but this ever-popular disclaimer is a poor substitute for satisfactory descriptive work. It should be recalled that Duncan and Duncan (1955) found  $D$  to be an acceptable index of racial segregation only because there appeared to be a "characteristic form for the segregation curves of most large American cities" (p. 214). While it is conventional to assume that the corresponding curves for sex segregation data are likewise invariant, such an assumption has not yet been verified in any rigorous way (but see Roos 1985, pp. 38–66; Gaskin 1979). At this relatively early point in the development of sex

<sup>4</sup> The comparisons based on margin-dependent measures will be especially misleading when carried out across countries or time periods that differ substantially in occupational composition or in rates of female labor force participation.

segregation research, we would thus agree with James and Taeuber (1985) that a "prudent analyst would construct and visually compare segregation curves" (p. 26; see also Brinton and Ngo [1991] for a related argument). We will be presenting such curves throughout the following analyses.

There is good reason to believe that the distinctive cultures and political histories of advanced industrial countries can live on in ways that affect the contours of their segregation profiles. Among scholars who study detailed patterns of sex segregation, it is now commonplace to cite the integration of Soviet women into medical and engineering positions (e.g., Roos 1985; Blekher 1979; Dodge 1971) or to note that pharmacy and dentistry are female-dominated occupations in Finland, Poland, and Hungary (e.g., Safilios-Rothschild 1976; see also Szelényi and Poster 1991). We will be carrying out analyses designed to reveal whether institutional differences of a more fundamental kind can also generate cross-national variation at the level of major occupational groupings. It has long been argued, for example, that Swedish women are disproportionately channeled into a state-supported service sector that is dominated by traditionally sex-typed jobs (see Ruggie 1984; Scriven 1984). By contrast, the well-known internal labor markets of Japan have the apparent effect of pushing women into blue-collar production work (see Brinton 1988; Kalleberg and Lincoln 1988; Clark 1979), while the traditional cultural and institutional environment of Switzerland encourages women to choose occupations that are compatible with their marginal forms of labor market attachment (see Charles and Buchmann 1994; Buchmann and Charles, *in press*). The structure of these nation-specific sex segregation systems has been reviewed in more detail elsewhere (see, e.g., Charles 1990); at this point, we merely wish to raise the possibility that some forms of cross-national variability may be revealed in *how* male and female work is separated, and not merely in the degree of such separation. If this is indeed the case, then the detailed contours of sex segregation cannot be adequately described with scalar indices, nor can satisfactory explanatory models be devised when these indices are applied.<sup>5</sup>

#### A GENERAL LOG-MULTIPLICATIVE APPROACH

We will proceed by fitting a series of association models that are consistent with the conventional practice of summarizing cross-national vari-

<sup>5</sup> The latter point may be obvious to those familiar with cross-national variability in gender stratification systems. We are merely commenting on the apparent disjuncture between our substantive understanding of such systems and conventional methods for describing them (see Brinton and Ngo 1993).

ability in a single parameter. The following model will serve as our starting point:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{(\phi_k Z_i v_j)}, \quad (3)$$

where  $i$  indexes sex,  $j$  indexes occupation,  $k$  indexes country,  $m_{ijk}$  is the expected frequency in cell  $(i, j, k)$ ,  $\alpha_k$  is the grand mean in the  $k$ th country,  $\beta_{ik}$  is the country-specific marginal effect for the  $i$ th gender,  $\gamma_{jk}$  is the country-specific marginal effect for the  $j$ th occupation,  $\phi_k$  is the multiplicative shift effect for the  $k$ th country,  $Z_i$  is an indicator variable for gender (i.e.,  $Z_1 = 0$  and  $Z_2 = 1$ ), and  $v_j$  is the scale value for the  $j$ th occupation.<sup>6</sup> The distinctive feature of this model is that it scales both occupations and countries *without* assuming any prior ranking (see Goodman 1979a, 1979b, 1981a; Clogg 1982; Xie 1992).<sup>7</sup> As indicated above, the sex-by-occupation association is expressed in a set of  $J$  column effects ( $v_j$ ), while any cross-national variability in sex segregation has to be absorbed by a set of  $K$  multiplicative shift effects ( $\phi_k$ ).<sup>8</sup> If this specification fits the data, it follows that the segregation profile is invariant and that  $\phi_k$  can be used to represent cross-national differences in the underlying "strength" of sex segregation. We will not be following the common practice of simply assuming that a scalar index suffices; instead, we have embedded this assumption in a testable model, with the viability of  $\phi_k$  as a segregation index thus resting on the usual criteria of model fit. If our multiplicative shift model fails to provide a satisfactory fit, we will have to modify it by permitting the shape of the sex segregation profile to vary across countries.

The full set of parameters in equation (3) cannot be uniquely estimated. In the following analyses, we will identify the column effects by con-

<sup>6</sup> We will be using the same indicator variable ( $Z_i$ ) for most of the models introduced in this article. The metric that we have used to scale gender is of course arbitrary and inconsequential; that is, not only are the column effects estimated here reproduced perfectly by all scalings in which  $Z_2 - Z_1 = 1$ , but they can be easily recovered (by multiplying through by  $c$ ) for all scalings in which  $Z_2 - Z_1 = c$ .

<sup>7</sup> We have found it convenient to parameterize the sex segregation profile in terms of column effects. However, given that our classification includes only two rows of data, this specification places no within-country constraints on the sex-by-occupation association. Although there are any number of parameterizations that might be adopted in this context, we think it is elegant and analytically revealing to choose one that represents both countries and occupations as scalable quantities. It should nonetheless be kept in mind that many of the log-multiplicative models introduced below could be rewritten as seemingly simpler log-linear models.

<sup>8</sup> Whereas prior analysts (e.g., Yamaguchi 1987) have parameterized shift effects in additive form, we have followed Xie (1992) in adopting a simple multiplicative specification (see also Clogg 1982; Becker and Clogg 1989; Fukumoto and Grusky 1992). The advantage in doing so is that the resulting model is invariant under all possible reorderings of the column and level categories (see Yamaguchi 1987, p. 484).

straining them to sum to zero and the marginal and multiplicative shift effects by constraining the parameters for the first row, column, or level to equal one:

$$\sum_{j=1}^J v_j = 0,$$

and (4)

$$\beta_{1k} = \gamma_{1k} = \phi_1 = 1.$$

The following closed-form results hold when the model represented by equation (3) fits perfectly:

$$\ln(\beta_{2k}) = 1/J \times \left[ \sum_{j=1}^J \ln(F_{jk}/M_{jk}) \right],$$

and

$$\begin{aligned} \phi_k v_j &= \ln(F_{jk}/M_{jk}) - \left[ 1/J \times \sum_{j=1}^J \ln(F_{jk}/M_{jk}) \right] \\ &= \ln(F_{jk}/M_{jk}) - \ln(\beta_{2k}). \end{aligned} \quad (5)$$

These results indicate that the main effect of gender for the  $k$ th country is merely the mean of the logged sex ratios, while the adjusted column effects ( $\phi_k v_j$ ) are simply occupation-specific departures from that mean.<sup>9</sup> If the model fails to fit perfectly, the estimates of  $\ln(\beta_{2k})$  and  $\phi_k v_j$  can of course be recovered by replacing  $F_{jk}$  and  $M_{jk}$  with their expected values.

#### THE SEX SEGREGATION DATA

We will apply this general model to an eight-nation data set collected by the International Labour Office (ILO). The ILO classifies labor force members into the following major occupations: (1) professional, technical, and related workers, (2) administrative and managerial workers, (3) clerical and related workers, (4) sales workers, (5) service workers, (6) production and related workers, and transport equipment operators and laborers, and (7) agricultural, animal husbandry, and forestry workers,

<sup>9</sup> The reader should be reminded that  $M_{jk} = m_{1jk}$  and that  $F_{jk} = m_{2jk}$ . We have maintained our earlier notation (see eqq. [1] and [2]) because it emphasizes the connection between conventional segregation indices and the log-multiplicative measures introduced here.



and fishermen and hunters.<sup>10</sup> We will eliminate the seventh category from the following analyses because of substantial cross-national discrepancies in the procedures for assigning women to agricultural labor (see ILO 1986, p. 3). The final sample counts for each country are presented in appendix A (for more details, see ILO [1985, 1986, 1987]).

The available evidence suggests that the rank ordering of countries on standard segregation indices remains roughly the same under both aggregated and disaggregated occupational classifications (see Jonung 1984; Charles 1990; Jacobs and Lim 1992). However, the standard indices may not fully reveal the potentially distorting effects of aggregation, and we will therefore be carrying out a series of supplementary analyses with disaggregated census data from the United States and Japan (see U.S. Bureau of the Census 1984, table 276, pp. 166–75; Statistics Bureau of Japan 1984, pp. 586–620). We have recoded the latter data into an aggregated version of the detailed 1968 International Standard Classification of Occupations (ISCO), with the aggregations being introduced whenever the original occupational codes could not sustain the more detailed classification.<sup>11</sup> The end result is the 45-category classification presented in appendix B.<sup>12</sup>

#### CONSTRUCTING A MARGIN-FREE INDEX OF SEX SEGREGATION

We will begin our analyses with a global test of cross-national variability in sex segregation. It could well be argued that a basic family resemblance in segregation regimes is generated by cross-nationally shared cultural, economic, and institutional forces. We are referring, in particular, to the worldwide diffusion of a family structure in which women have primary responsibility for childrearing, cooking, and maintaining the home. According to some theorists, women in such families should not only have a diminished incentive to invest in work-related human capital (given that investments must often be repaid over a shorter work life) but should also prefer occupations in which wage depreciation is minimized during

<sup>10</sup> These major categories are based on the 1968 International Standard Classification of Occupations (ISCO), but for convenience we shall refer to them as "ILO categories."

<sup>11</sup> We are indebted to Harry Ganzeboom for his recoding of the Japanese data into the ISCO classification.

<sup>12</sup> Although the following analyses will be based exclusively on sex segregation tables, it would be equally instructive to analyze other forms of segregation data with our general multiplicative approach (e.g., racial occupational segregation). Indeed, whenever there is a one-to-one correspondence between the categories being compared (e.g., occupations), the analyst will usually wish to apply measures that are margin free.

TABLE 1  
LOG-LINEAR AND LOG-MULTIPLICATIVE ASSOCIATION MODELS  
APPLIED TO EIGHT-NATION DATA SET

Model or Contrast	$L^2$	$df$	$L^2/L^2_1$	$\Delta$
Model:				
1. Conditional independence ( $O \times N + S \times N$ ) .....	7,121,442	40	100.0	18.1
2. Constant sex segregation ( $O \times N + S \times N + C$ ) .....	789,207	35	11.1	4.1
3. Multiplicative shift effect ( $O \times N + S \times N + C + A \times N$ ) .	527,952	28	7.4	2.4
4. Segregation profiles ( $O \times N + S \times N + C \times P$ ) .....	126,389	25	1.8	1.3
Contrast:				
1. Total variability (model 2) .....	789,207	35	100.0	
2. Explained variability under model 3 (model 3 vs model 2) .....	261,255	7	33.1	
3. Explained variability under model 4 (model 4 vs. model 2) .....	662,818	10	84.0	

NOTE.— $O$  = occupation,  $N$  = country;  $S$  = sex;  $C$  = column effects;  $A$  = global row-by-column association parameter;  $P$  = tripartite partition of countries. The index of dissimilarity ( $\Delta$ ) can be interpreted as the percentage of cases that would have to be reallocated to bring the observed and expected values into perfect correspondence.

the many employment interruptions that their family obligations require (see Polachek and Siebert 1993; Polachek 1979, 1981; cf. England et al. 1988). The foregoing arguments suggest that men in all countries should dominate jobs that entail a substantial commitment to the labor market, whereas women should be disproportionately concentrated in jobs that neither require large amounts of human capital nor strongly penalize employment interruptions. Although the traditional family structure may therefore have a powerful standardizing effect on gender stratification systems, we would also cite the cultural diffusion of occupational gender labels and gender-based stereotypes as potential forces for convergence.<sup>13</sup>

The overall effect of such forces can be assessed by fixing  $\phi_k$  at one (for all  $k$ ) and thereby forcing all countries to share the same segregation profile. As shown in table 1 (see model 2), we find a substantial amount of cross-national similarity in our eight-nation data set, with the model of constant segregation misallocating only 4.1% of the cases and account-

<sup>13</sup> The occupational categories deployed here are so highly aggregated that one would expect substantial within-category variability on such dimensions as human capital requirements, wage depreciation, and the like. This heterogeneity greatly reduces our leverage in adjudicating between competing arguments about the sources or causes of occupational segregation.

ing for 88.9% of the association under the model of conditional independence. The column effects under model 2 take on a largely expected form; namely, women are underrepresented in managerial and production occupations, and they are overrepresented in clerical, sales, and service occupations (see fig. 2)<sup>14</sup> This evidence of cross-national convergence might be seen as the analogue within the field of sex segregation to the Featherman-Jones-Hauser finding that "the genotypical pattern of mobility . . . in industrial societies with a market economy and a nuclear family system is basically the same" (Hauser and Featherman 1977, p. 16; Featherman, Jones, and Hauser 1975).

At the same time, it is important to explore the cross-national variability in the data, since the model of constant sex segregation clearly fails to fit. The first step in doing so is to ask whether a more limited form of convergence might hold; that is, we can entertain the provisional hypothesis that the underlying segregation profile takes on the same basic shape in each country, while the overall *degree* of segregation is cross-nationally variable. This hypothesis leads us to our full log-multiplicative model (see eq. [3]) in which the peaks and valleys of a generic segregation profile are compressed or expanded in accord with a simple shift effect.<sup>15</sup> The latter model does indeed reveal considerable cross-national variation; as indicated in figure 3, the peaks and valleys for the Swedish curve are 2.7 times more extreme than those for the relatively flat Japanese curve, whereas the results for the six remaining countries fall somewhere between these two poles. However, we would not want to take these estimates too seriously, since our baseline model does not adequately account for the cross-national variability in the data. Under a standard decomposition of the test statistic (for model 2 of table 1), only 33.1% of the total variability can be explained with multiplicative shift effects, while the remaining variability arises from cross-national differences in the structure of the segregation profile itself (see table 1, contrast 2).

The latter result indicates that  $\phi_k$  cannot satisfactorily represent the structure of cross-national variability. Moreover,  $\phi_k$  is not necessarily the best of all possible scalar indices, since it improperly assumes a common segregation profile. If we insist on defining a scalar index in such circumstances, we would be well advised to base it on a model that fits the

<sup>14</sup> The pattern of sex segregation revealed here may not be entirely surprising, but it is nonetheless difficult to reconcile it with the simple hypothesis that women should be disproportionately represented in occupations that require only limited investments in human capital. The latter hypothesis is neither consistent with the negative scale value for production workers nor with the gender neutral scale value for professionals.

<sup>15</sup> The log-multiplicative models presented here were estimated with modified GLIM programs that were based on and inspired by similar programs prepared by Mark P. Becker and Yu Xie.

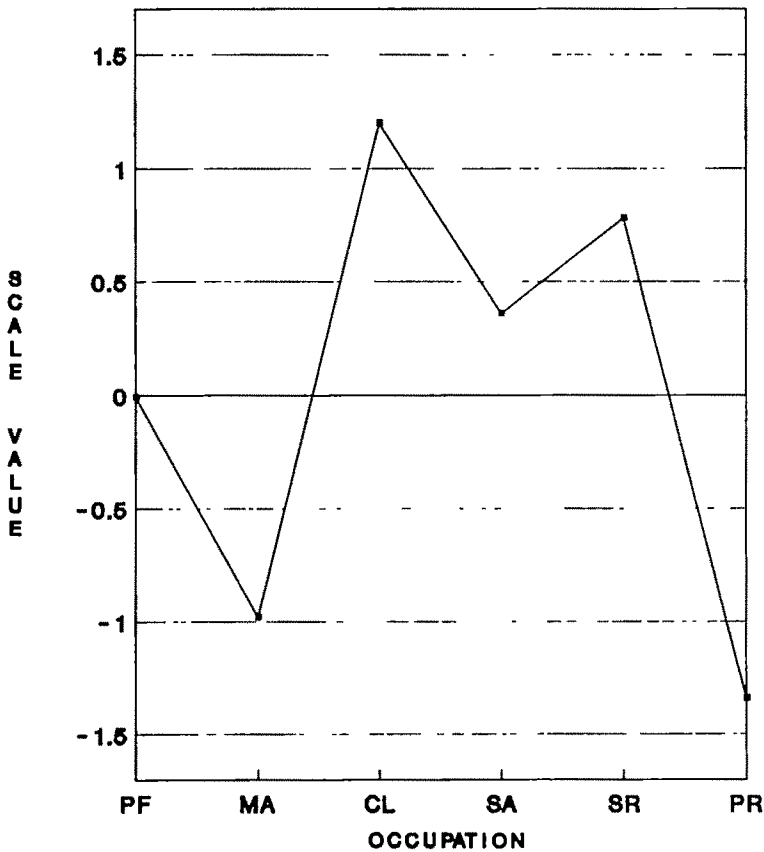


FIG. 2.—A pooled profile of sex segregation. The parameter estimates are taken from model 2 (table 1). In the present graph and all following ones, the positive scale values indicate female overrepresentation and the negative scale values indicate male overrepresentation. We have used the following abbreviations for the ILO occupations: PF = professional; MA = manager; CL = clerical; SA = sales; SR = service; PR = production.

three-way association between occupation, sex, and country.<sup>16</sup> This model can be represented as follows:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{(Z_i v_{jk})}, \quad (6)$$

<sup>16</sup> It is no easy task to choose among scalar indices when the multiplicative shift model fails to fit. In this context, one must either (1) distort the data at the point of estimating the model, or (2) apply a model that fits perfectly and then summarize the many parameters of that model in a single scalar index (thereby introducing "distortions" of a different sort). Although some information will necessarily be lost with either approach, we advocate the latter one because the resulting index (i.e.,  $A$ ) is sensitive to all departures from perfect integration rather than merely those which emerge under a particular representation of the common segregation profile.

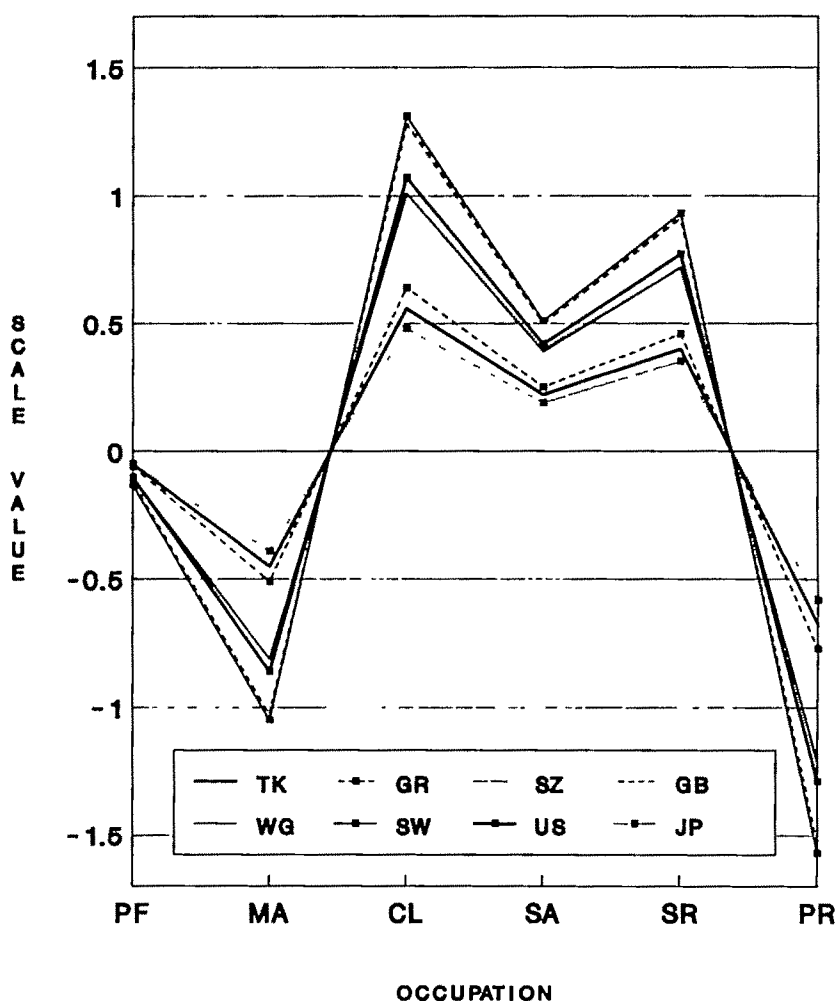


FIG. 3.—Country-specific levels of sex segregation. The parameter estimates are taken from model 3 (table 1). We have used the following abbreviations for the ILO occupations: PF = professional; MA = manager; CL = clerical; SA = sales; SR = service; PR = production. We have also abbreviated the names of countries: TK = Turkey; GR = Greece; SZ = Switzerland; GB = Great Britain; WG = West Germany; SW = Sweden; US = United States; JP = Japan.

where each of the letters and subscripts is defined as before.<sup>17</sup> The column effects under this saturated model can be used to calculate the expected occupation-specific deviation from perfect integration:

$$\begin{aligned} A_k &= \exp \left( 1/J \times \sum_{j=1}^J v_{jk}^2 \right)^{1/2} \\ &= \exp \left( 1/J \times \sum_{j=1}^J \left\{ \ln(F_{jk}/M_{jk}) - \left[ 1/J \times \sum_{j=1}^J \ln(F_{jk}/M_{jk}) \right] \right\}^2 \right)^{1/2}. \end{aligned} \quad (7)$$

This new index of association is closely related to  $\phi_k$ .<sup>18</sup> Indeed, when our multiplicative shift model fits perfectly,  $\phi_k$  is completely governed by the size of  $A$ . If, for example,  $A_s$  and  $A_t$  denote the value of  $A$  in any two (arbitrarily chosen) countries, then the following result holds:

$$\ln(A_s) = \phi_s/\phi_t \times \ln(A_t). \quad (8)$$

We can therefore conclude that  $\ln(A_s)/\ln(A_t)$  will equal  $\phi_s/\phi_t$  whenever equation (3) correctly characterizes the observed data.

Unlike the size-standardized index of dissimilarity, our new index is invariant under multiplicative transformations of the sex ratio. When  $F_j$  is multiplied by the factor  $c$  for all  $j$ ,  $A$  can be reexpressed as follows:

$$\begin{aligned} A'_k &= \exp \left( 1/J \times \sum_{j=1}^J \left\{ \ln(cF_{jk}/M_{jk}) - \left[ 1/J \times \sum_{j=1}^J \ln(cF_{jk}/M_{jk}) \right] \right\}^2 \right)^{1/2} \\ &= \exp \left( 1/J \times \sum_{j=1}^J \left\{ \ln(c) + \ln(F_{jk}/M_{jk}) \right. \right. \\ &\quad \left. \left. - \left[ 1/J \times \sum_{j=1}^J \ln(c) + \ln(F_{jk}/M_{jk}) \right] \right\}^2 \right)^{1/2} \\ &= \exp \left( 1/J \times \sum_{j=1}^J \left\{ \ln(F_{jk}/M_{jk}) - \left[ 1/J \times \sum_{j=1}^J \ln(F_{jk}/M_{jk}) \right] \right\}^2 \right)^{1/2} \\ &= A_k. \end{aligned} \quad (9)$$

<sup>17</sup> The marginal effects for this model are identified as before (see eq. [4]), whereas the column effects are now constrained to sum to zero within each country.

<sup>18</sup> In an earlier draft of this article, we used the absolute value function in defining  $A$  (see also Charles 1992). The present definition has the virtue of making the relationship between  $A_k$  and  $\phi_k$  more transparent (see eq. [8]).

TABLE 2  
SCALAR MEASURES OF OCCUPATIONAL  
SEX SEGREGATION APPLIED TO  
EIGHT-NATION DATA SET

COUNTRY	SEGREGATION INDEX		
	<i>D</i>	<i>D<sub>s</sub></i>	<i>A</i>
Switzerland ... ..	39.9	41.6	3.53
Sweden .. . . . .	41.2	41.9	3.07
Great Britain.....	44.4	41.4	2.77
Turkey .... . . . .	40.5	46.2	2.64
Japan .... . . . .	24.1	30.1	2.56
Germany .... . . .	38.9	34.2	2.41
United States .....	36.6	28.9	2.41
Greece . . . . . . .	30.2	27.8	1.90

NOTE.—*D* = index of dissimilarity, *D<sub>s</sub>* = size-standardized index of dissimilarity, *A* = global association index under saturated model. The values of *A* are calculated from the saturated model (see eq. [6]).

As shown here, the constant factor “cancels out” under the algebra of natural logarithms, and *A'* thus reduces to *A*. Moreover, *A* is also unchanged when the *j*th occupational margin is multiplied by a constant, because  $(cF_{jk}/cM_{jk}) = (F_{jk}/M_{jk})$  for any arbitrary *c*. It follows from this result that *A* can be safely used to compare countries with different occupational distributions.

The values of *D*, *D<sub>s</sub>*, and *A* for our eight-nation data are presented in table 2. This table indicates that the rank ordering of countries under our margin-free index differs in nontrivial ways from what prevails under *D* or *D<sub>s</sub>*. For example, Switzerland registers the highest segregation level under *A*, while it occupies a middle position under *D*. We also find that Japan and Turkey are outliers under *D* and *D<sub>s</sub>*, but not under *A*. The conventional segregation scores for Japan are usually regarded as surprisingly low (see Brinton and Ngo 1991); however, the present results indicate that the intrinsic association in Japan is quite strong, with our new index implying that males or females are overrepresented in the average Japanese occupation by a factor of 2.56.

#### CONSTRUCTING SEX SEGREGATION PROFILES

The test statistics presented in table 1 (see contrast 2) make it clear that any scalar index will conceal cross-national variability in the underlying segregation profile. As indicated earlier, we do have some hypotheses

TABLE 3

OCCUPATION-SPECIFIC SEX SEGREGATION PARAMETERS FOR EIGHT-NATION DATA SET

COUNTRY	COLUMN EFFECT					
	Professional	Manager	Clerical	Sales	Service	Production
Turkey.....	1.28	-.74	1.41	-1.01	-.40	-.55
Greece.....	.61	-.82	.84	-.19	.33	-.77
Sweden....	.22	-1.29	1.50	-.10	1.16	-1.50
Japan.....	.55	-1.90	.87	-.03	.74	-.23
United States.....	.01	-.53	1.45	.01	.49	-1.43
Switzerland..	.19	-2.20	.80	.93	1.39	-1.12
Great Britain.....	-.22	-1.02	1.27	.58	.92	-1.52
Germany.....	.10	-1.14	.91	.69	.68	-1.24

NOTE.—The parameter estimates are taken from the saturated model (see eq. [6]). The positive estimates indicate female overrepresentation, and the negative estimates indicate male overrepresentation. Estimates that are large and positive are shown in *italic*.

about the contours of these cross-national differences, but for our present purposes it is useful to proceed inductively by grouping the unconstrained column effects into distinct segregation profiles. The results from this exercise are presented in table 3 and figure 4.

The column effects in table 3 provide limited evidence of cross-national parallelism in occupational sex typing. In all eight countries, we find that women are concentrated in the vast "middle class" of clerical work, while men consistently dominate managerial and production occupations. It would be a mistake, however, to gloss over the rather substantial cross-national differences in the *degree* of sex segregation within these categories. For example, Japanese males are overrepresented in production work by a factor of only 1.26, whereas the corresponding sex ratios for Great Britain and Sweden are as high as 4.57 and 4.48 ( $\exp[.23] = 1.26$ ;  $\exp[1.52] = 4.57$ ;  $\exp[1.50] = 4.48$ ).<sup>19</sup> We would further note that the American managerial sector is highly integrated by current international standards; indeed, American males are only 1.70 times more likely than their female counterparts to be managers, whereas the corresponding sex ratios for Sweden, Japan, and Switzerland are two to five times stronger ( $\exp[0.53] = 1.70$ ;  $\exp[1.29] = 3.63$ ;  $\exp[1.90] = 6.69$ ;  $\exp[2.20] = 9.03$ ).<sup>20</sup> These results are clearly consistent with some of the

<sup>19</sup> We have reversed the sign of the coefficients in table 3, since we are now reporting the male-to-female sex ratio.

<sup>20</sup> The contrast between the Swedish and American sex ratios may reflect the different ways in which these countries have accommodated demands for reduced gender stratification. As argued elsewhere, legalistic guarantees of equal opportunity (e.g., "affirmative action" policies) have historically played a lesser role in Swedish-style corporatist systems, since interest groups in such systems tend to be co-opted with more



country-specific institutional characteristics that we briefly discussed in our introductory comments.

The sex ratios for professional, sales, and service occupations provide additional evidence that the segregation systems of industrial societies have not yet converged to a common pattern. In each of these occupations, we find cross-national inconsistencies in the direction of occupational sex typing; it is these types of "sign shifts" in the column effects that account for the poor fit of our baseline specification (see table 1, model 3). The most extreme examples of such shifts can be found in the segregation profiles for Turkey and Greece. As shown in figure 4, sales occupations are male dominated under profile A (see esp. Turkey), while the same occupations are gender neutral under profile B and female dominated under profile C. Obversely, we find that the professional sector is female dominated under profile A but gender neutral (or more nearly so) under profiles B and C.<sup>21</sup> Although results of this general sort could always be attributed to cross-national differences in the composition of the major occupations (see below for details), we suspect that true sociocultural forces can account for some of the sign shifts. It is hardly surprising, for example, that Turkish women have been driven away from sales and service work, since traditional Islamic law strictly regulates the role of women in commercial transactions (see Arat 1989).

The foregoing graphs thus indicate that Turkey and Greece have a top-heavy pattern of segregation. We can further distinguish between two types of bottom-heavy profiles: the first one has a bimodal cast (see profile B), whereas the second one is less jagged in the interior regions (see profile C). The resulting typology can be usefully formalized by estimating a model that fits a single segregation curve for each profile. As shown in table 1, the latter model fits quite well; in fact, our profile-specific specification correctly allocates 98.7% of the respondents (see table 1, model 4), and it accounts for approximately 84% of the total cross-national variability (see table 1, contrast 3). The remaining variability is partly attributable to differences in the degree of sex segregation within each profile.

---

tangible concessions (e.g., state-financed child care and guaranteed pregnancy leaves). The pluralistic and legalistic political structure of the United States has been translated into affirmative action programs rather than state-financed concessions of the Swedish sort (see Lovenduski 1986; Silén 1988; Gelb 1989; Charles 1990, 1992).

<sup>21</sup> It could well be argued that the sex segregation regime of Japan is more properly classified under profile A. In this regard, we should emphasize that our profiles obviously cannot be viewed as definitive, but rather are merely illustrative representations of the types of log-multiplicative analyses that might be carried out.

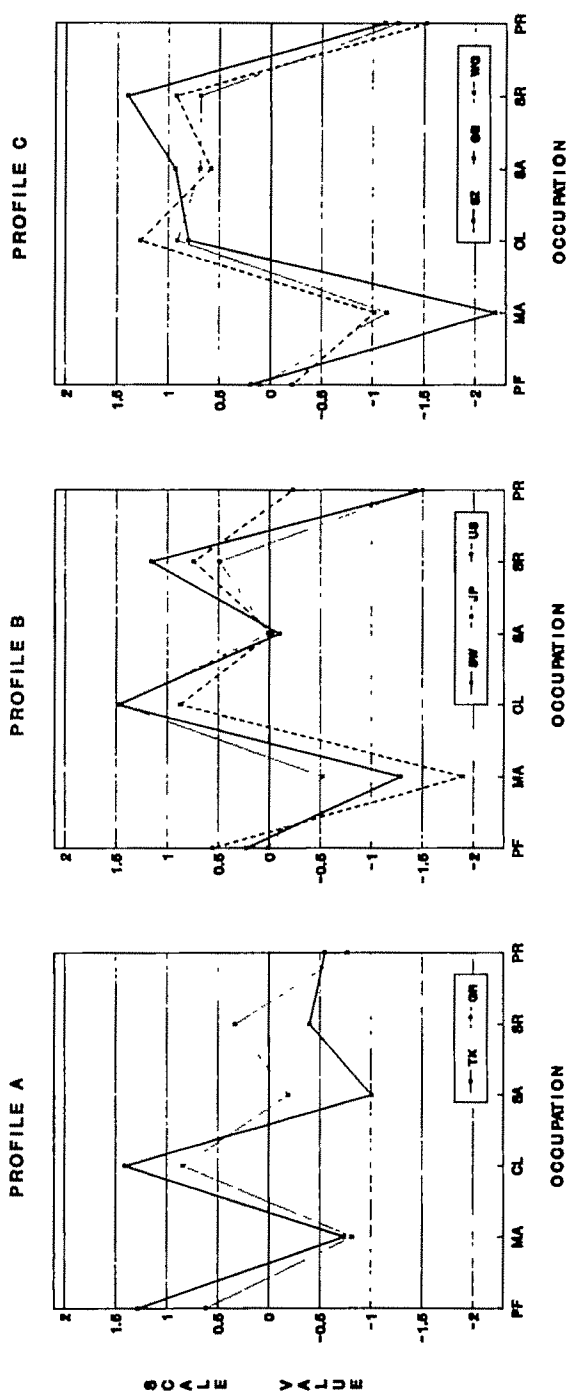


FIG. 4.—Three profiles of occupational sex segregation in eight-nation data set. The parameter estimates are taken from the saturated model (see eq. [6]). We have used the following abbreviations for the ILO occupations: PF = professional; MA = manager; CL = clerical; SA = sales; SR = service; PR = production. We have also abbreviated the names of countries: TK = Turkey; GR = Greece; SZ = Switzerland; GB = Great Britain; WG = West Germany; SW = Sweden; US = United States; JP = Japan.

# MODELS FOR DISAGGREGATED DATA

Up to now, our analyses have been based exclusively on the aggregated ILO classifications, but clearly the same types of models and methods could be readily applied to disaggregated data. The purpose of the present section is to introduce a series of multiplicative models that provide new insights into the effects of aggregation and disaggregation on segregation statistics. We will address the following questions in turn:

1. How much of the total sex-by-occupation association is lost by aggregating the data into six ILO categories? Can we capture the most important cross-national differences in sex segregation with highly aggregated data?
2. Are the aggregated ILO profiles (see fig. 4) distorted by cross-national differences in the mixture of detailed occupations found within each major category? Does a cross-nationally common curve emerge when these compositional biases are purged from the data?
3. Can we characterize the detailed segregation profiles with a multiplicative shift parameter? Does the rank-ordering of countries under this shift parameter mimic the corresponding rank-ordering for aggregate data (see table 2)?

As we noted earlier, this set of supplementary analyses will have to be carried out with our two-nation sample (see app. B), since the ILO classifications are only published in highly aggregated form. The results presented in this section should therefore be seen as largely illustrative in intent.

We will begin by asking whether the aggregated ILO classification conceals a substantial amount of sex-by-occupation association. This question can be addressed by fitting the following model for each country:

$$m_{ij} = \alpha\beta_i\gamma_j e^{(z_i v_c^*)}, \quad (10)$$

where

$$\begin{array}{ll} c = 1 & \text{if } 1 \leq j \leq 13, \\ c = 3 & \text{if } 16 \leq j \leq 22, \\ c = 5 & \text{if } 26 \leq j \leq 31, \end{array} \quad \begin{array}{ll} c = 2 & \text{if } 14 \leq j \leq 15, \\ c = 4 & \text{if } 23 \leq j \leq 25, \\ c = 6 & \text{if } 32 \leq j \leq 45. \end{array}$$

The above set of "side constraints" on  $v_c^*$  forces the column effects to be equal within each of the six major ILO categories.<sup>22</sup> We have not fitted a corresponding set of column effects for detailed occupations; conse-

<sup>22</sup> This model can be identified by fixing  $\beta_1$  and  $\gamma_1$  at "1" and by requiring the six column effects ( $v_c^*$ ) to sum to "0."

TABLE 4  
COMPONENTS OF TOTAL ASSOCIATION IN DISAGGREGATED DATA  
FROM UNITED STATES AND JAPAN

Model	$L^2$	$df$	$L^2/L^1$	$\Delta$
United States:				
Independence ( $O + S$ ) ..	6,902,112	44	100 0	25.6
Major occupation effects ( $O + S + G$ )...	3,219,930	39	46.7	13 1
Japan:				
Independence ( $O + S$ ) .	2,481,293	44	100 0	20 1
Major occupation effects ( $O + S + G$ )..	1,663,678	39	67 0	12 8

NOTE.— $O$  = detailed occupation,  $S$  = sex,  $G$  = major occupation column effects

quently, the model implies that the segregation regime is homogeneous across occupations after the ILO effects are parsed out, with the associated test statistic thus indexing the extent to which the data violate this implication (see Goodman 1981*b*; Bishop, Fienberg, and Holland 1975, pp. 126–30; Featherman and Hauser 1978, pp. 180–84; Allison 1980). As shown in table 4, 46.7% of the association in the American table is lost by aggregating the data into six ILO categories, while a full 67.0% of the Japanese association is lost under the same aggregation. In interpreting these results, some scholars might prefer to emphasize that a relatively large component of association can be explained with only 5  $df$ , whereas others might point out that most of the sex-by-occupation association is generated within the six ILO categories. The latter type of description tends to be preferred by contemporary segregation researchers; that is, the prevailing view seems to be that the standard occupational groupings (e.g., the ILO categories) are unacceptably heterogeneous and that researchers should therefore attempt to ratchet segregation analyses down to the lowest possible level (see, e.g., Cain 1984; Reskin and Roos 1987, p. 11; Sokoloff 1987, p. 62).<sup>23</sup> We see nothing in the results of table 4 that is inconsistent with such an interpretation.

It is quite another matter to ask whether aggregated data will typically misinform us about the structure of cross-national variability in sex segregation. The sources of this variability can be specified with the following model:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{(Z_i v_j + Z_i v_{ik}^2)}, \quad (11)$$

<sup>23</sup> This viewpoint is typically informed by the pathbreaking work of Bielby and Baron (1984, 1986).

TABLE 5  
LOG-LINEAR AND LOG-MULTIPLICATIVE ASSOCIATION MODELS  
APPLIED TO TWO-NATION DATA SET

Model	$L^2$	$df$	$L^2/L^1$	$\Delta$
Constant sex segregation ( $O \times N + S \times N + C$ ).....	730,843	44	100.0	4.9
Additive shift in means ( $O \times N + S \times N + C + G \times N$ ).....	220,520	39	30.2	2.0
Hybrid model ( $O \times N + S \times N + C + A \times N + G \times N$ )...	219,457	38	30.0	2.0

NOTE.— $O$  = detailed occupation;  $N$  = country;  $S$  = sex;  $C$  = detailed occupation column effects;  $A$  = global row-by-column association parameter;  $G$  = major occupation column effects.

where  $\nu_j$  refers to the unconstrained column effects and  $\nu_c^*$  refers to the ILO column effects that are generated by the six sets of equality constraints implied by equation (10).<sup>24</sup> Under the specification indicated here, the only outlet for cross-national variability is the shift effects defined at the level of major categories, since the scale values for the detailed occupations are constrained to be the same across countries. It turns out that this type of constraint does not lead us too far astray; in fact, the model represented by equation (11) misclassifies only 2% of the respondents, and it accounts for nearly 70% of the total cross-national variability (see table 5, "Additive shift in means"). We can thus conclude that the most important institutional differences in sex aggregation are expressed at the level of major occupational groupings. Although the ILO categories clearly conceal a substantial amount of sex-by-occupation association (see table 4), it would appear that this residual association takes on a relatively similar form in each country.

The latter result does not give us full license to proceed with an aggregate analysis. After all, the occupational composition of the ILO categories may well differ across countries, and it is therefore possible that some of the cross-national variability observed in figure 4 is artifactual (see Sokoloff [1987, pp. 63–66] for a related argument). We can eliminate this form of "compositional bias" by fitting the following model:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{(Z_i \nu_{jk} + Z_i \nu_{ik}^*)}, \quad (12)$$

<sup>24</sup> The identifying restrictions deployed here are directly analogous to those represented by eq. (4). That is, the set of six ILO column effects for the second country ( $\nu_{c2}^*$ ) are constrained to sum to zero, and the pooled set of 45 microlevel column effects ( $\nu_j$ ) are likewise constrained to sum to zero. The ILO column effects for the first country ( $\nu_{c1}^*$ ) are all fixed at one.

where  $v_j$  and  $v_c^*$  are defined as before.<sup>25</sup> The distinctive feature of this model is that it parameterizes the structure of segregation at multiple levels, thereby making it possible to estimate the net residue of segregation at the aggregate level after first purging the data of lower-order compositional effects (see Stier and Grusky [1990] for a related model; also, for a pathbreaking discussion of purging, see Clogg and Eliason [1988] and Clogg, Shockey, and Eliason [1990]). We have graphed the relevant results from this model in figure 5, with the dotted lines denoting the purged ILO column effects, and the solid lines denoting the unpurged effects calculated from the aggregated data.<sup>26</sup> The overall picture that emerges from this figure is only partially reassuring; to be sure, some of the original cross-national differences persist in unchanged form (i.e., PF and PR), but the remaining ones are either strengthened or weakened. While this finding indicates that the contours of our segregation profiles (see fig. 4) may be somewhat distorted, it is worth noting that the purged column effects in figure 5 continue to vary across the two nations (see esp. MA, SA, and PR). We must therefore reject our provisional hypothesis that such biases were concealing a cross-nationally common profile.

In some circumstances, analysts of sex segregation may also wish to characterize the structure of the detailed segregation curve, and their attention will thus turn to some form of multiplicative shift model. As may be recalled, we were forced to reject this highly restrictive specification for the aggregated ILO data (see table 1, contrast 2), but it is still possible that the segregation profile for detailed occupations will take on a more structured form. This possibility can be addressed by fitting the following model:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{[\phi_k(Z_i, v_j) + Z_i v_{ck}^*]}, \quad (13)$$

where  $\phi_k$  is a multiplicative shift parameter governing the relative height of the peaks and valleys in the detailed segregation profile.<sup>27</sup> As indicated

<sup>25</sup> The model presented here will always fit the data perfectly. The microlevel column effects ( $v_{jk}$ ) were identified by being forced to sum to zero within the six major occupational categories of each country, whereas the ILO column effects ( $v_{ck}^*$ ) were identified by being forced to sum to zero within each country. When these restrictions are imposed, the microlevel column effects account for 39 df in each country (i.e.,  $45 - 6 = 39$ ), and the ILO column effects account for the remaining 5 df in each country (i.e.,  $6 - 1 = 5$ ).

<sup>26</sup> The unpurged column effects in fig. 5 differ from the corresponding effects in fig. 4, since the two sets of estimates are based on different data.

<sup>27</sup> The shift effects ( $\phi_k$ ) for this model were identified by fixing  $\phi_1$  at one, and the column effects ( $v_j$  and  $v_{ck}^*$ ) were identified by imposing the restrictions specified in n. 24. For didactic purposes, we have presented two sets of means ( $v_{c1}^*$  and  $v_{c2}^*$ ) in fig. 6, but it should be kept in mind that one of these sets is implied by the microlevel column effects and does not, therefore, convey any additional information.

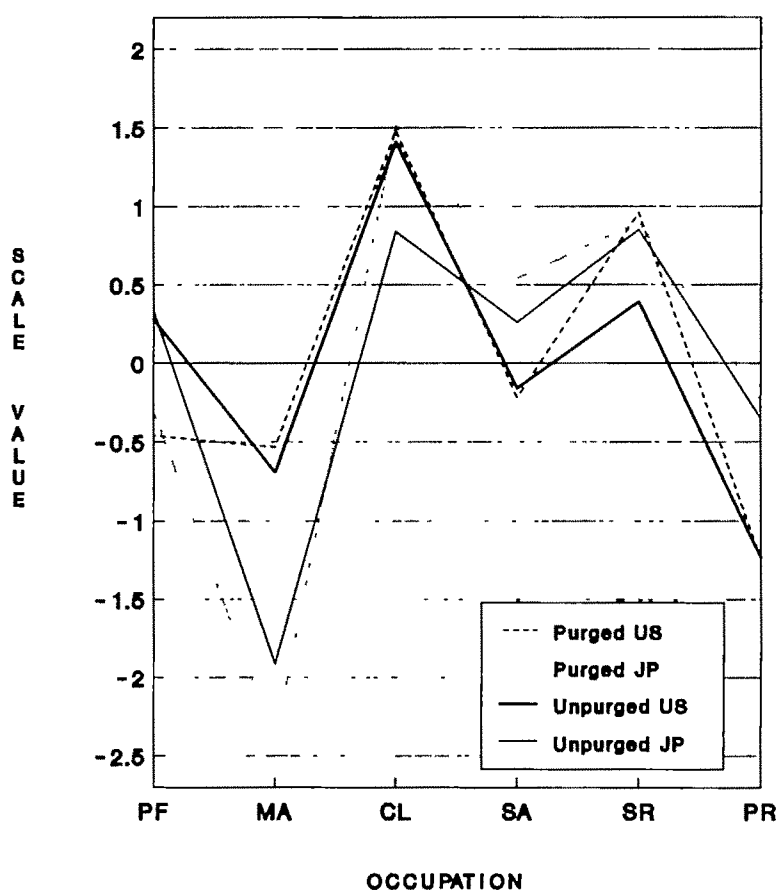


FIG. 5.—Purged and unpurged means for two-nation data. The purged scale values are drawn from the saturated model represented in eq. (12). We have used the following abbreviations for the ILO occupations: PF = professional; MA = manager; CL = clerical; SA = sales; SR = service; PR = production. We have also used abbreviated names for Japan (JP) and the United States (US).

in figure 6, this new model permits the purged ILO means ( $\nu_c^*$ ) to freely vary across countries, but it constrains the lower-order column effects ( $\nu_j$ ) to be a multiplicative function of  $\phi_k$  (see table 5 for the relevant fit statistics). In the present case, the estimated ratio of  $\phi_1$  to  $\phi_2$  equals 1.047, and we can thereby conclude that the dispersion around the ILO means is approximately 4.7% greater in Japan than in the United States. It follows that Japan has a higher segregation index at two levels of analysis; that is, not only are the Japanese ILO categories more segre-

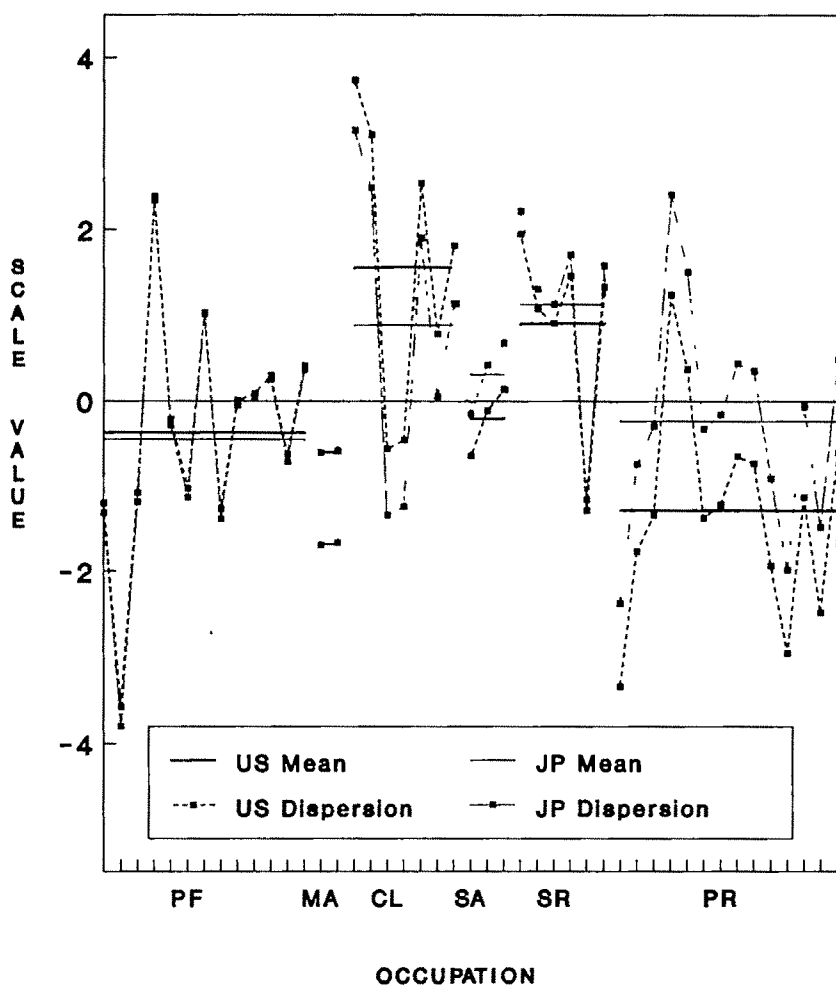


FIG. 6.—Hybrid multiplicative model for two-nation data. The parameter estimates are taken from the hybrid model (table 5). We have used the following abbreviations for the ILO occupations: PF = professional; MA = manager; CL = clerical; SA = sales; SR = service; PR = production. We have also used abbreviated names for Japan (JP) and the United States (US).



gated than the corresponding American categories, but so too are the detailed occupations within these categories.<sup>28</sup>

#### THE HIERARCHICAL STRUCTURE OF SEX SEGREGATION

The models that we have presented up to now all treat the occupational structure as a simple nominal variable. Although most of the standard segregation indices (e.g.,  $D$ ,  $D_i$ ) also rest on a nominal level of measurement, there has been a recent resurgence of work based on alternative indices that require some form of ordinal or continuous scaling (see Brinton and Ngo 1991; Stafford and Fossett 1989; Fossett, Galle, and Kelly 1986; White 1983). This work has been motivated, at least in part, by the so-called checkerboard problem that White (1983) first discussed in the context of residential segregation research. As was noted by White (1983, pp. 1010–11), the value of  $D$  will be unaffected by the spatial arrangement of the underlying census tracts (i.e., their “checkerboard layout”), since the relevant calculations for  $D$  are based on the internal composition of the tracts rather than the distances between them (see also Duncan and Duncan 1955, p. 215; Taeuber and Taeuber 1965, p. 205). The implication, of course, is that  $D$  will fail to register “tract-level” desegregation; for example, if a large ghetto were broken up by scattering the constituent tracts throughout the city, the value of  $D$  would necessarily remain unchanged.<sup>29</sup>

The same type of problem arises in a simpler (one-dimensional) form when occupational data are analyzed. In the latter context, one might wish to calculate the “social distance” between the male and female distributions, with the relevant metric typically being some form of prestige or socioeconomic scale. It is clearly inappropriate to use  $D$  for such purposes, since it was designed to measure “nominal differentiation rather than inequality” (Fossett et al. 1986, p. 423). Indeed, just as  $D$  cannot detect the residential desegregation that occurs when an all-black tract is moved to a neighborhood formerly dominated by whites, so too it cannot detect the sex desegregation that occurs when an all-female occupation is moved to a “socioeconomic region” formerly dominated by males (see also Fossett and South 1983; Stafford and Fossett 1989, p.

<sup>28</sup> If the multiplicative shift effects are allowed to be category specific, we find that in two cases (i.e., clerical and production occupations) the segregation index is actually larger in the United States than in Japan. This elaborated model fits relatively well ( $L^2 = 193,005$ ), but it still accounts for only 12.5% of the total variability in the detailed segregation profile ( $1 - [193,005/220,520] = .125$ ). We are thus well advised to treat the estimates from our multiplicative shift models with some caution.

<sup>29</sup> The obvious irony here is that the checkerboard problem can be solved (in ad hoc fashion) by resorting to *less* detailed levels of measurement.

179; Brinton and Ngo 1991). This deficiency has motivated some segregation scholars to define and deploy alternative indices that take into account the location of census tracts, school districts, or occupations in physical or social space. It should come as no surprise that these revised indices are often direct modifications of  $D$  (e.g., Brinton and Ngo 1991).

Unless the indices so proposed are functions of the relevant cross-product ratios, they will again be margin-dependent and therefore flawed for comparative purposes. We can secure a margin-free measure by fitting a scaled association model of the following kind:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{\psi_k(Z_i, T_{jk})}, \quad (14)$$

where  $\psi_k$  and  $T_{jk}$  are direct analogues to  $\phi_k$  and  $v_{jk}$ , and the remaining parameters retain their original meaning. This model differs from our baseline specification in equation (3) because  $T_{jk}$  refers to a priori values rather than freely estimated ones.<sup>30</sup> We thus end up with a hybrid specification that stands somewhere between the association models of Haberman (1974) and those of Hout (1984). As indicated in equation (14), the row categories in our data are scaled with the standard unit scores of a linear-by-linear interaction model (see Haberman 1974; Duncan 1979; Goodman 1979a), whereas the column categories are scaled with external scores of the kind deployed by Hout (1984, 1988), Hauser (1984), and others (Hout and Jackson 1986; Szelenyi 1988).

We have illustrated this simple approach by estimating  $\psi_k$  conditional on the prestige scores from SIOPS (Standard International Occupational Prestige Scale).<sup>31</sup> Under our specification,  $T_{j1}$  equals  $T_{j2}$  for all  $j$ ; this constraint holds because we have scaled the occupations in each country with the same prestige scores. The fit statistics in table 6 indicate that our model misallocates 23.7% of the respondents and accounts for only 1.2% of the total sex-by-occupation association (see the scaled association model). It would thus appear that SIOPS cannot adequately account for the underlying structure of sex segregation; if anything, the graphs in figure 7 suggest that the column effects take on a curvilinear form, with the inflection point in both countries occurring between 45 and 50 prestige points. This poor performance is of course consistent with some of our prior results. The graphs in figure 5, for example, suggested that the ILO

<sup>30</sup> The multiplicative shift parameter indexes the overall sex-by-occupation association conditional on the scores so assigned.

<sup>31</sup> In almost all cases, the occupations in our 45-category classification correspond to a "minor group" in ISCO, and we could therefore directly apply the published ISCO version of the standard scale (see Treiman 1977, app. A). The remaining scores were estimated by averaging across the SIOPS values for all of the "unit occupations" contained within a given category (see app. B below for a listing of the final scores).

TABLE 6  
MODELS OF SCALED ASSOCIATION FOR TWO-NATION DATA

Model	$L^2$	$df$	$L^2/L^1$	$\Delta$
Conditional independence ( $O \times N + S \times N$ ) .....	9,383,405	88	100.0	23.7
Scaled association ( $O \times N + S \times N + P \times N$ ) .....	9,274,708	86	98.8	23.7
Purged scaled association ( $O \times N + S \times N + P \times G \times N$ )	3,667,937	66	39.1	10.1

NOTE.— $O$  = detailed occupation;  $N$  = country;  $S$  = sex;  $P$  = product term (row scores  $\times$  column scores);  $G$  = major occupation.

column effects do not follow a simple prestige gradient, whereas our additional tests in table 4 indicated that nearly half of the association in the disaggregated tables is generated at the ILO level. These two findings imply that a simple association model was doomed from the start.

The expected prestige gradient may nonetheless emerge after the ILO effects are purged from the data. It is commonly argued, for example, that aggregate analyses of sex segregation are misleading because males tend to secure the most desirable occupations within each of the major categories conventionally deployed. This expectation can be tested, albeit only partially, by fitting a model of the following kind:

$$m_{ijk} = \alpha_k \beta_{ik} \gamma_{jk} e^{[\psi_{ik}^*(Z_i, T_i) + Z_i v_{ik}^*]}, \quad (15)$$

where  $v_{ik}^*$  refers to the ILO column effects in the  $k$ th country (see eq. [10] for the relevant equality constraints),  $\psi_{ik}^*$  refers to the corresponding country-specific association parameters estimated within each of the six ILO major categories, and the remaining letters and subscripts are defined as before.<sup>32</sup> As might be expected, this revised specification fits relatively well; the results in table 6 reveal that only 39.1% of the association remains unexplained when  $v_{ik}^*$  and  $\psi_{ik}^*$  are included as additional terms. While the fit statistics are much improved under this specification, we still find that the resulting prestige gradient does not take on the conventionally expected shape (see table 7). In fact, males enjoy an advantage over females in only three ILO categories (professional, sales, and service), whereas the reverse association prevails in the two re-

<sup>32</sup> This model can be estimated by applying our original association model (see eq. [14]) to the six subtables formed by disaggregating across the ILO categories. If the data are analyzed as a single array (as represented in eq. [15]), then the ILO column effects ( $v_{ik}^*$ ) can be identified by being forced to sum to zero within each country.

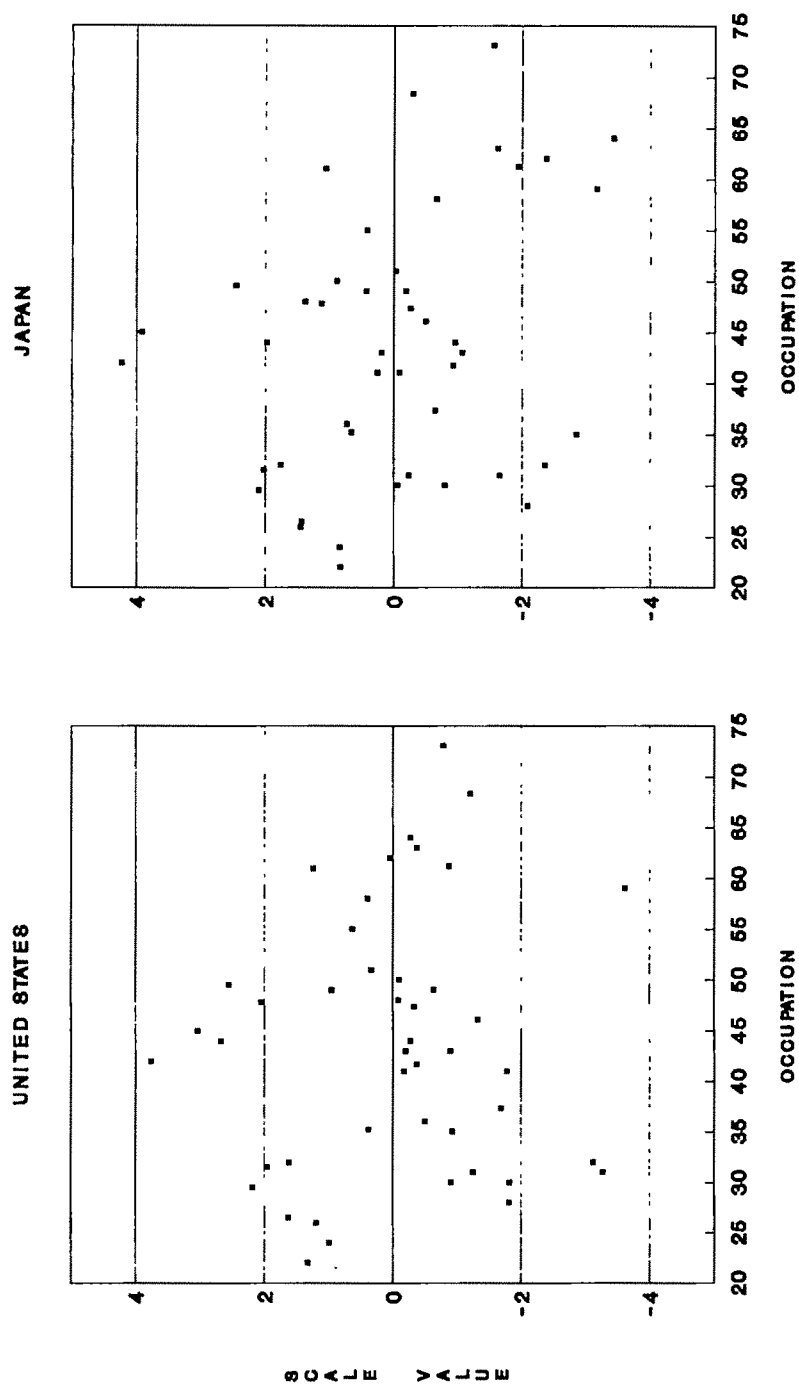


Fig. 7.—Scatter plots of column effects by prestige for the United States and Japan. The scale values are taken from the saturated model. The horizontal axis indexes the SIOPS scores for the detailed occupations in our 45-category classification.

TABLE 7  
PARAMETERS OF SCALED ASSOCIATION FOR  
MAJOR OCCUPATIONAL CATEGORIES

OCCUPATION	ASSOCIATION INDEX	
	United States	Japan
Professional... ..	-.113	-.118
Clerical .....	.217	.112
Sales .....	-.044	-.040
Service. ....	-.213	-.184
Production .....	.010	.021

NOTE.—The parameter estimates are taken from model 3, table 6. The positive estimates indicate female overrepresentation, and the negative estimates indicate male overrepresentation. We have omitted the managerial category because it contains only two occupations.

maining cases (clerical and production). If our SIOPS scores were a perfect measure of the general desirability of jobs (see Goldthorpe and Hope 1974), these results would be partially inconsistent with the hypothesis that "occupational composition [is] the result of a matching process in which the top-ranked workers get the most attractive jobs" (Reskin and Roos 1990, p. 307; see also Strober 1984).<sup>33</sup>

## CONCLUSIONS

We began this article by noting that the conventional segregation indices are dependent on the marginal distributions in a sex-by-occupation array (see fig. 1). It was this deficiency that motivated Gibbs (1965), Gross (1968), and other segregation scholars to modify  $D$  by standardizing for "differences among occupational categories with regard to their share of the labor force" (Gibbs 1965, p. 163). The resulting size-standardized index ( $D_s$ ) has now become the measure of choice among contemporary scholars who seek to compare segregation regimes across time or space (e.g., Brinton and Ngo 1991, 1993; Presser and Kishor 1991; Williams 1979; Jacobs 1989a, 1989b; Jacobs and Lim 1992). It should be empha-

<sup>33</sup> It should be emphasized, of course, that a queuing hypothesis of this sort cannot be adequately addressed with the present data. After all, SIOPS is clearly not an exhaustive measure of job desirability (see Jencks et al. 1988), nor is our 45-category classification sufficiently fine grained to reveal the structure of male advantage at the most detailed level of analysis (see Bielby and Baron 1984). We would further note that a queuing model cannot be tested in any convincing way without controlling for heterogeneity in worker qualifications (e.g., education, work experience, and training).

sized that the transition to  $D_s$  occurred without great fanfare; indeed,  $D_s$  became popular well after the methodological debates of the 1960s had run their course, and it was therefore shielded from the rigorous vetting that earlier indices had undergone (see Peach 1975). This is not to say that the standardization proposed by Gibbs (1965) failed in its stated objective to eliminate the distorting effects of the occupational structure. The size-standardized index does indeed live up to its billing; however, the cost of standardizing in this fashion is that  $D_s$  is no longer scale invariant (see James and Taeuber 1985), and hence researchers replacing  $D$  with  $D_s$  are merely exchanging one form of marginal dependence for another.

If we wish to construct a margin-free index, we have no choice but to resort to measures that are functions of cross-product ratios. The centerpiece of our approach has been a simple association model that permits the peaks and valleys of the segregation profile to be compressed or expanded in accord with a scalar shift effect. We have thus rejected the conventional practice of assuming that a scalar index is adequate to the task; instead, we have argued that this assumption should be embedded in a testable model, with the viability of a scalar approach resting on the usual criteria of model fit. As it turns out, only one-third of the total cross-national variability can be explained with a multiplicative shift effect, while the remaining variability must be attributed to heterogeneity in the segregation curves themselves (see table 1, contrast 2). Although the latter form of heterogeneity cannot be fully captured with a scalar index, we have nonetheless defined a margin-free measure ( $A$ ) that is based on the particular segregation profile prevailing within each country. We have also used elaborated versions of our general model to examine the hierarchical structure of segregation, to identify the dominant "segregation profiles" in industrial countries, and to parse out the net residue of segregation at multiple levels of analysis.

This modeling framework may prove to be useful in future analyses of occupational sex segregation. At the same time, the conventional indices have so far shown remarkable staying power, and it would thus be presumptuous of us to suppose that the forces of inertia and entrenched interest will be quickly or easily overcome. Among the various arguments that are likely to be raised against our analyses, the following are perhaps the most obvious ones:

*Issues of aggregation.*—By the usual standards of sex segregation research, our occupational categories are highly aggregated, and some skeptics may therefore find our results to be unconvincing or even misleading. In this regard, we would emphasize that our analyses were completed at the aggregate level for reasons of convenience and data avail-

ability, and not because of any intrinsic limitations of our modeling framework.<sup>34</sup> We would nonetheless contend that patterns of sex segregation at the major occupational level are of considerable interest because they signify correspondingly major differences in socioeconomic rewards and conditions. In our continuing efforts to ferret out segregation at the most detailed level possible, we ought not to forget that much of the occupational heterogeneity in life chances, work conditions, and consumption practices is likely located at the level of major categories. While the analytic returns to disaggregating occupations may therefore be limited, this is not to gainsay the equally important point that research carried out at the major occupational level is subject to compositional biases that can and should be purged by the methods that we introduced above (see eq. [12]).

*Conceptualizing sex segregation.*—It is also important to address the fallback argument that log-multiplicative measures are not sufficiently faithful to traditional conceptualizations of sex segregation. If sex segregation is defined to be whatever  $D$  or  $D_i$  measure, then of course such arguments hold in a nominal sense. We would suggest, however, that many researchers *have* adopted an implicit conceptualization of segregation that is distinct from these common operationalizations (e.g., Williams 1979; Blau and Hendricks 1979; England 1981; Bridges 1982; Handl 1984; Tienda and Ortiz 1987; Beller 1984; Bianchi and Rytina 1986; Abrahamson and Sigelman 1987; Jacobs 1989a, 1989b; Jacobs and Lim 1992; Presser and Kishor 1991; Brinton and Ngo 1991, 1993). Indeed, in treating column effects as the fundamental parameters of sex segregation, we have merely operationalized the long-standing assumption that such parameters are properly independent of both the occupational structure and the rate of female labor force participation. The development of  $D_i$  was seemingly motivated by a similar objective, yet it failed to fully realize the implicit conceptualization that underlies it.<sup>35</sup>

*Issues of endogeneity.*—In making the prior point, we do not mean to suggest that traditional conceptualizations of segregation are necessarily desirable, nor that any clear consensus on conceptual matters has now been reached. However, we would argue that scholars who reject margin-free measures of sex segregation should not likewise reject all forms

<sup>34</sup> If a cross-classification is disaggregated to the point of being extremely sparse, then asymptotic results are less safely presumed and sampling zeroes may frequently appear. Although some of the models introduced here cannot be directly estimated when sampling zeros are present, the methods for analyzing such arrays are relatively well developed (see, e.g., Clogg and Eliason 1987).

<sup>35</sup> While  $D$  also confounds marginal and interaction effects, its great virtue is that it does so in a way that has a pleasing interpretation. The same cannot be said for  $D_i$ , and we would therefore recommend abandoning the latter measure altogether.

of multiplicative modeling, since many competing representations of segregation can be elegantly operationalized in the context of simple log-linear or log-multiplicative specifications. This includes, for example, parameterizations in which the rate of female labor force participation is not exogenously determined, but rather is generated by the conjunction of occupational marginal effects and occupation-specific gender ratios. The latter formulation is hardly radical or innovative; it underlies, in fact, all arguments to the effect that the secular trend in female labor force participation has been generated by the exogenous growth of female-typed occupations (see, e.g., Oppenheimer 1970). The following model is suggested by such arguments:

$$m_{ij} = \alpha \gamma_j \delta_{ij}, \quad (16)$$

where  $\gamma_1 = 1$  and  $\delta_{11} = \delta_{12} = \dots = \delta_{1J} = 1$ . The segregation parameters under this specification (i.e.,  $\delta_{ij}$ ) are a mixture of the marginal ( $\beta_j$ ) and interaction ( $\nu_j$ ) effects estimated under our saturated multiplicative model (see eq. [6]). In a single sex-by-occupation array, these two specifications are perforce equivalent, since both fit the data perfectly. However, when a third dimension is introduced (e.g., time), it becomes possible to determine which of these specifications might be construed as structural and thereby preferred (see Duncan 1975; also, for a relevant application, see Grusky and Hauser [1984]). The model of equation (16) could be elaborated in various other ways, but for our present purposes it should suffice to emphasize that multiplicative models are consistent with a wide range of segregation parameterizations, not all of which are margin free.

The appeal of a multiplicative framework thus rests in large part with its analytic flexibility. We would like to conclude by reviewing, if only briefly, some potentially useful modifications and extensions of our preferred models. Among the more straightforward extensions, we would include (1) models that incorporate additional scalable variables representing further contexts in which segregation processes are nested (e.g., age, period, cohort, industry, and firm size), (2) models that account for cross-context variability in segregation by conditioning on a priori scores that characterize the contexts,<sup>36</sup> and (3) models of racial segregation that freely scale both the occupational categories and the racial categories

<sup>36</sup> The requisite models here would be similar in structure to those represented in eqq. (14) and (15). However, rather than applying external scores to the occupational categories, we would now be applying such scores to the categories indexing countries. The contextual variables represented in this fashion can be constrained to exert global effects on all occupations or to exert particularized effects that are specific to certain occupations or combinations of occupations (Charles 1992, also, see Grusky and Hauser 1984; Hauser and Grusky 1988)



represented in a race-by-occupation array (e.g., whites, African-Americans, Hispanics, and Native Americans). It would also be useful to revise our preferred models in more ambitious and far-reaching ways by explicitly incorporating individual-level covariates (e.g., education). In doing so, cross-national differences in the endowments or human capital of men and women could be effectively controlled, with the result being a purged version of *A* that indexes the residual variability generated in the labor market itself (see Yamaguchi [1983] for a related model). The latter approach would provide a useful bridge between descriptive measures of sex segregation and standard explanatory models of sex discrimination.

## APPENDIX A

TABLE A1

OBSERVED COUNTS IN CROSS-CLASSIFICATION OF SEX  
BY OCCUPATION IN EIGHT COUNTRIES

COUNTRY AND SEX	OCCUPATION					
	Professional	Managerial	Clerical	Sales	Service	Production
Turkey:						
Male.....	580,983	148,629	437,380	709,755	833,713	3,675,554
Female.. ..	244,868	8,310	209,217	30,132	65,342	247,207
Greece:						
Male.. ..	20,029	5,296	18,311	21,688	18,061	93,755
Female.....	12,440	789	14,310	6,055	8,498	14,744
Switzerland:						
Male.....	290,252	69,673	294,564	110,684	115,560	913,171
Female....	177,659	3,921	330,847	141,404	235,065	151,017
Great Britain:						
Male.....	2,497,820	1,787,150	1,011,550	571,510	1,004,260	6,796,060
Female .....	1,639,970	524,560	2,933,700	832,620	2,058,480	1,208,680
Germany:						
Male.....	9,528	3,336	8,048	4,236	5,444	30,752
Female .....	6,496	660	12,408	5,232	6,652	5,488
Sweden:						
Male.....	6,006	869	1,089	2,057	1,628	11,407
Female .....	7,183	231	4,675	1,793	4,994	2,453
United States:						
Male. ....	69,007	62,813	28,041	53,110	48,578	211,708
Female .....	65,988	34,936	112,601	50,809	74,607	47,986
Japan:						
Male.....	28,710	19,503	44,649	51,381	21,780	147,312
Female....	23,661	1,386	50,787	23,661	21,780	55,440

NOTE.—We have deflated the population estimates provided by the ILO into sample counts (see Charles 1990).

## APPENDIX B

TABLE B1

OBSERVED COUNTS IN DETAILED CROSS-CLASSIFICATION OF SEX  
BY OCCUPATION IN UNITED STATES AND JAPAN

OCCUPATION	COUNTRY AND SEX			
	U S. Males	U S. Females	Japanese Males	Japanese Females
<b>Professional:</b>				
Researcher and engineer (61.2)...	642,634	157,536	245,666	13,703
Air and ship officer (59.0).....	21,202	336	13,285	218
Medical professional (68.3)....	122,463	21,672	45,432	13,315
Other medical workers (49.5).....	94,147	704,711	36,709	166,849
Accountant (62.0).....	119,046	73,397	7,917	286
Jurist (73.0).....	91,240	24,445	8,026	661
Teacher (61.0)....	286,338	582,908	163,535	184,642
Worker in religion (46.0)....	54,687	8,543	18,233	4,288
Author and journalist (58.0).....	31,034	27,140	16,930	3,394
Painter and sculptor (51.0).....	61,094	50,303	23,481	8,889
Musician and performer (48.0)...	33,406	18,005	10,645	16,559
Athlete (49.0).....	7,644	2,388	4,188	1,345
Other professional (55.0).....	164,548	182,423	33,225	19,639
<b>Manager:</b>				
Government official (64.0).....	48,413	21,634	28,107	355
Business manager (63.0).....	1,048,860	425,148	463,655	35,827
<b>Clerical:</b>				
Stenographer and typist (42.0)...	6,003	151,256	598	16,000
Key punch operator (45.0).....	5,695	69,320	507	9,885
Transport conductor (41.7).....	145,761	58,894	34,524	5,328
Mail clerk (44.0).....	119,115	53,483	34,778	5,195
Telephone operator (44.0).....	6,482	55,062	7,280	20,494
Machine operator (49.0).....	43,163	65,504	10,145	6,048
Other clerical (47.8).....	515,164	2,340,363	806,349	965,553
<b>Sales:</b>				
Sales manager (47.3).....	261,998	110,705	247,146	73,908
Agent and broker (50.0).....	170,983	90,689	56,924	53,705
Other sales (35.2).....	620,865	531,291	629,740	474,062
<b>Service:</b>				
Maid and housekeeping (29.5) ..	40,017	208,478	8,298	26,331
Cook and food service (26.0).....	312,831	607,134	210,059	346,838
Launderer (22.0).....	17,604	38,756	16,725	14,985
Personal service (32.0)....	33,383	98,422	37,629	84,949
Protective service (35.0).....	585,171	135,701	151,266	3,411
Other service (26.5).....	56,372	168,474	37,135	60,555
<b>Production:</b>				
Extractive (32.0).....	66,537	1,726	14,019	517
Metal and plastic worker (37.3)	537,146	58,543	498,643	101,921
Wood worker (31.0).....	45,608	7,668	73,493	22,621
Textile worker (31.5).....	56,147	230,627	94,081	277,981

TABLE B1 (Continued)

OCCUPATION	COUNTRY AND SEX			
	U S. Males	U.S. Females	Japanese Males	Japanese Females
Shoemaker (24.0) ... ..	7,945	12,582	6,291	5,689
Assembler and repairer (43.0) ....	792,613	189,202	288,013	38,506
Electrical worker (41.0) . . . . .	296,981	29,446	217,322	109,276
Jeweler (43.0).....	4,865	2,350	3,941	1,857
Printer (41.0) . . . . .	73,689	36,438	57,624	20,540
Painter (30.0) . . . . .	102,942	9,844	61,258	10,740
Construction worker (31.0).... ..	734,172	16,250	555,789	41,301
Machine operator (30.0) . . . . .	611,655	144,905	170,537	62,691
Transport equipment (28.0) . . . . .	564,560	54,343	471,832	22,686
Other production (36.0) . . . . .	1,177,269	416,955	356,730	289,370

NOTE.—The entries in parentheses are the estimated Standard International Occupational Prestige scores (Treiman 1977)

## REFERENCES

- Abrahamson, Mark, and Lee Sigelman. 1987. "Occupational Sex Segregation in Metropolitan Areas." *American Sociological Review* 52:588-97.
- Allison, Paul. 1980. "Analyzing Collapsed Contingency Tables without Actually Collapsing." *American Sociological Review* 45:123-30.
- Altham, Patricia M. E. 1970. "The Measurement of Association of Rows and Columns for an  $r \times s$  Contingency Table." *Journal of the Royal Statistical Society, ser. B*, 32:63-73.
- Arat, Yesim. 1989. *The Patriarchal Paradox: Women Politicians in Turkey*. London: Associated University Press.
- Baron, James N., Brian S. Mittman, and Andrew E. Newman. 1991. "Targets of Opportunity. Organizational and Environmental Determinants of Gender Integration within the California Civil Service, 1979-1985." *American Journal of Sociology* 96:1362-1401.
- Becker, Mark P., and Clifford C. Clogg. 1989. "Analysis of Sets of Two-Way Contingency Tables Using Association Models." *Journal of the American Statistical Association* 84:142-51.
- Bell, Wendell. 1954. "A Probability Model of the Measurement of Ecological Segregation." *Social Forces* 32:357-64.
- Beller, Andrea H. 1984. "Trends in Occupational Segregation by Sex and Race, 1960-1981." Pp. 11-26 in *Sex Segregation in the Workplace*, edited by Barbara F. Reskin. Washington, D.C.: National Academy.
- Bianchi, Suzanne M., and Nancy Rytina. 1986. "The Decline in Occupational Sex Segregation during the 1970s: Census and CPS Comparisons." *Demography* 23: 79-86.
- Bielby, William T., and James N. Baron. 1984. "A Women's Place Is with Other Women: Sex Segregation within Organizations." Pp. 27-55 in *Sex Segregation in the Workplace*, edited by Barbara F. Reskin. Washington, D.C.: National Academy.
- . 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91:759-99.

- Bishop, Yvonne M., Stephen E. Fienberg, and Paul W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT Press.
- Blau, Francine D., and Wallace E. Hendricks. 1979. "Occupational Segregation by Sex: Trends and Prospects." *Journal of Human Resources* 14:197-210.
- Blekher, Feiga. 1979. *The Soviet Woman in the Family and in Society*. New York: Wiley.
- Bridges, William P. 1982. "The Sexual Segregation of Occupations: Theories of Labor Stratification in Industry." *American Journal of Sociology* 88:270-95.
- Brinton, Mary C. 1988. "The Social-Institutional Bases of Gender Stratification: Japan as an Illustrative Case." *American Journal of Sociology* 94:300-334.
- Brinton, Mary C., and Hang-Yue Ngo. 1991. "Occupational Sex Segregation in Comparative Perspective." Manuscript. University of Chicago, Department of Sociology.
- . 1993. "Age and Sex in the Occupational Structure: A United States-Japan Comparison." *Sociological Forum* 8:93-111.
- Buchmann, Marlis, and Maria Charles. In press. "Organizational and Institutional Factors in the Process of Gender Stratification: Comparing Social Arrangements in Six European Countries." *International Journal of Sociology*.
- Butler, Richard J. 1987. "New Indices of Segregation." *Economics Letters* 24: 359-62.
- Cain, Pamela Stone. 1984. "Commentary." Pp. 87-90 in *Sex Segregation in the Workplace*, edited by Barbara F. Reskin. Washington, D.C.: National Academy.
- Charles, Maria. 1990. *Occupational Sex Segregation: A Log-Linear Analysis of Patterns in 25 Industrial Countries*. Ph.D. dissertation. Stanford University, Department of Sociology.
- . 1992. "Cross-national Variation in Occupational Sex Segregation." *American Sociological Review* 57:483-502.
- Charles, Maria, and Marlis Buchmann. 1994. "Assessing Micro-Level Explanations of Occupational Sex Segregation: Human Capital Development and Labor Market Opportunities in Switzerland." *Swiss Journal of Sociology* vol. 20, no. 3.
- Clark, Rodney. 1979. *The Japanese Company*. New Haven, Conn.: Yale University Press.
- Clogg, Clifford C. 1982. "Using Association Models in Sociological Research: Some Examples." *American Journal of Sociology* 88:114-34.
- Clogg, Clifford C., and Scott R. Eliason. 1987. "Some Common Problems in Log-Linear Analysis." *Sociological Methods and Research* 16:8-44.
- . 1988. "A Flexible Procedure for Adjusting Rates and Proportions, Including Statistical Methods for Group Comparisons." *American Sociological Review* 53: 267-83.
- Clogg, Clifford C., James W. Shockey, and Scott R. Eliason. 1990. "A General Statistical Framework for Adjustment of Rates." *Sociological Methods and Research* 19:156-95.
- Coulter, Philip B. 1989. *Measuring Inequality*. Boulder, Colo.: Westview.
- Dodge, Norton. 1971. "Women in the Soviet Economy." Pp. 207-23 in *The Professional Woman*, edited by Athena Theodore. Cambridge, Mass.: Schenkman.
- Duncan, Otis D. 1975. *Introduction to Structural Equation Models*. New York: Academic Press.
- . 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 4:793-803.
- . 1984. Foreword. Pp. ix-xii in *The Analysis of Cross-Classified Data Having Ordered Categories*, by Leo A. Goodman. Cambridge, Mass.: Harvard University Press.
- Duncan, Otis D., and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indices." *American Sociological Review* 20:210-17.

- England, Paula. 1981. "Assessing Trends in Occupational Sex Segregation, 1900-1976." Pp. 273-95 in *Sociological Perspectives on Labor Markets*, edited by Ivar Berg. New York: Academic Press.
- England, Paula, George Farkas, Barbara S. Kilbourne, and Thomas Dou. 1988. "Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects." *American Sociological Review* 53:544-58.
- Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.
- Featherman, David L., F. Lancaster Jones, and Robert M. Hauser. 1975. "Assumptions of Mobility Research in the United States: The Case of Occupational Status." *Social Science Research* 4:329-60.
- Fossett, Mark A., and Scott J. South. 1983. "The Measurement of Intergroup Income Inequality: A Conceptual Review." *Social Forces* 61:855-71.
- Fossett, Mark A., Omer R. Galle, and William R. Kelly. 1986. "Racial Occupational Inequality, 1940-1980: National and Regional Trends." *American Sociological Review* 51:421-29.
- Fuchs, Victor R. 1975. "A Note on Sex Segregation in Professional Occupations." *Explorations in Economic Research* 2:105-11.
- Fukumoto, Ivan, and David B. Grusky. 1992. "Social Mobility and Class Structure in Early-Industrial France." Pp. 40-67 in *Building European Society: Occupational Change and Social Mobility in Europe, 1840-1940*, edited by Andrew Miles and David Vincent. Manchester: Manchester University Press.
- Gaskin, Katherine A. 1979. *Occupational Differentiation by Sex: An International Comparison*. Ph.D. dissertation. University of Michigan, Department of Sociology.
- Gelb, Joyce. 1989. *Feminism and Politics: A Comparative Perspective*. Berkeley and Los Angeles: University of California Press.
- Gibbs, Jack P. 1965. "Occupational Differentiation of Negroes and Whites in the United States." *Social Forces* 44:159-65.
- Goldthorpe, John W., and Keith Hope. 1974. *The Social Grading of Occupations: A New Approach and Scale*. Oxford: Clarendon.
- Goodman, Leo A. 1979a. "Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-Classification Tables." *American Journal of Sociology* 84:804-19.
- . 1979b. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 70: 755-68.
- . 1981a. "Association Models and Canonical Correlation in the Analysis of Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 76:320-34.
- . 1981b. "Criteria for Determining Whether Certain Categories in a Cross-Classification Table Should Be Combined, with Special Reference to Occupational Categories in an Occupational Mobility Table." *American Journal of Sociology* 87:612-50.
- . 1991. "Measures, Models, and Graphical Displays in the Analysis of Cross-Classified Data." *Journal of the American Statistical Association* 86:1085-1111.
- Gross, Edward. 1968. "Plus ça change . . . ? The Sexual Structure of Occupations over Time." *Social Problems* 16:198-208.
- Grusky, David B., and Robert M. Hauser. 1984. "Comparative Mobility Revisited: Models of Convergence and Divergence in 16 Countries." *American Sociological Review* 49:19-38.
- Haberman, Shelby J. 1974. "Log-Linear Models for Frequency Tables with Ordered Classifications." *Biometrics* 30:589-600.
- Hakim, Catherine. 1992. "Explaining Trends in Occupational Segregation: The Mea-

- surement, Causes, and Consequences of the Sexual Division of Labour." *European Sociological Review* 8:127-52.
- Handl, Johann. 1984. "Chancengleichheit und Segregation: Ein Vorschlag zur Messung ungleicher Chancenstrukturen und ihrer zeitlichen Entwicklung." *Zeitschrift der Soziologie* 13:328-45.
- Hauser, Robert M. 1978. "A Structural Model of the Mobility Table." *Social Forces* 56:919-53.
- . 1984. "Vertical Mobility in Great Britain, France, and Sweden." *Acta Sociologica* 27:87-110.
- Hauser, Robert M., and David L. Featherman. 1977. *The Process of Stratification*. New York: Academic Press.
- Hauser, Robert M., and David B. Grusky. 1988. "Cross-National Variation in Occupational Distributions, Relative Mobility Chances, and Intergenerational Shifts in Occupational Distributions." *American Sociological Review* 53:723-48.
- Hout, Michael. 1984. "Status, Autonomy, and Training in Occupational Mobility." *American Journal of Sociology* 89:1379-1409.
- . 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358-1400.
- Hout, Michael, and John A. Jackson. 1986. "Dimensions of Occupational Mobility in the Republic of Ireland." *European Sociological Review* 2:114-37.
- ILO (International Labor Office). 1985. *Yearbook of Labour Statistics*. Geneva: ILO.
- . 1986. *Yearbook of Labour Statistics*. Geneva: ILO.
- . 1987. *Yearbook of Labour Statistics*. Geneva: ILO.
- Jacobs, Jerry A. 1989a. "Long-Term Trends in Occupational Segregation by Sex." *American Journal of Sociology* 95:160-73.
- . 1989b. *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, Calif.: Stanford University Press.
- Jacobs, Jerry A., and Suet T. Lim. 1992. "Trends in Occupational and Industrial Sex Segregation in 56 Countries, 1960-1980." *Work and Occupations* 19:450-86.
- James, David R., and Karl E. Taeuber. 1985. "Measures of Segregation." Pp. 1-32 in *Sociological Methodology 1985*, edited by Nancy Tuma. New York: Jossey-Bass.
- Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. "What Is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93:1322-57.
- Jonung, Christiana. 1984. "Patterns of Occupational Segregation by Sex in the Labor Market." Pp. 44-63 in *Sex Discrimination and Equal Opportunity*, edited by Günther Schmid and Renate Weitzel. Hampshire: Gower.
- Kalleberg, Arne L., and James R. Lincoln. 1988. "The Structure of Earnings Inequality in the United States and Japan." *American Journal of Sociology* 94:S121-S153.
- Lieberson, Stanley. 1969. "Measuring Population Diversity." *American Sociological Review* 34:850-62.
- . 1980. *A Piece of the Pie*. Berkeley and Los Angeles: University of California Press.
- . 1981. "An Asymmetrical Approach to Segregation." Pp. 61-82 in *Ethnic Segregation in Cities*, edited by Ceri Peach, Vaughan Robinson, and Susan Smith. Athens: University of Georgia Press.
- Lovenduski, Joni. 1986. *Women and European Politics: Contemporary Feminism and Public Policy*. Amherst: University of Massachusetts Press.
- Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67:281-315.
- Oppenheimer, Valerie. 1970. *The Female Labor Force in the United States: Demographic and Economic Factors Governing Its Growth and Changing Composition*. Berkeley: University of California Press.
- Peach, Ceri. 1975. *Urban Residential Segregation*. London: Longman.

- Polachek, Solomon. 1979. "Occupational Segregation among Women: Theory, Evidence, and a Prognosis." Pp. 137-57 in *Women in the Labor Market*, edited by Cynthia Lloyd, Emily Andrews, and Curtis Gilroy. New York: Columbia University Press.
- . 1981. "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics* 63:60-69.
- Polachek, Solomon, and W. Stanley Siebert. 1993. *The Economics of Earnings*. Cambridge: Cambridge University Press.
- Presser, Harriet B., and Sunita Kishor. 1991. "Economic Development and Occupational Sex Segregation in Puerto Rico, 1950-80." *Population and Development Review* 17:53-85.
- Reskin, Barbara F. 1993. "Sex Segregation in the Workplace." Pp. 241-70 in *Annual Review of Sociology*, vol. 19. Edited by Judith Blake and John Hagan. Palo Alto, Calif.: Annual Reviews.
- Reskin, Barbara F., and Patricia A. Roos. 1987. "Status Hierarchies and Sex Segregation." Pp. 3-21 in *Ingredients for Women's Employment Policy*, edited by Christine Bose and Glenna Spitze. New York: SUNY Press.
- . 1990. *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Rogoff, Natalie. 1953. *Recent Trends in Occupational Mobility*. Glencoe, Ill.: Free Press.
- Roos, Patricia A. 1985. *Gender and Work: A Comparative Analysis of Industrial Societies*. New York: SUNY Press.
- Ruggie, Mary. 1984. *The State and Working Women: A Comparative Study of Britain and Sweden*. Princeton, N.J.: Princeton University Press.
- Safilios-Rothschild, Constantina. 1976. "Dual Linkages between the Occupational and Family Systems: A Macrosociological Analysis." Pp. 51-60 in *Women and the Workplace. The Implications of Occupational Segregation*, edited by Martha Blaxall and Barbara Reagan. Chicago: University of Chicago Press.
- Scriven, Jeannie. 1984. "Women at Work in Sweden." Pp. 153-181 in *Working Women: An International Survey*, edited by Marilyn J. Davidson and Cary L. Cooper. New York: Wiley.
- Semyonov, Moshe. 1980. "The Social Context of Women's Labor Force Participation: A Comparative Analysis." *American Journal of Sociology* 86:534-50.
- Semyonov, Moshe, and Richard Ira Scott. 1983. "Industrial Shifts, Female Employment and Occupational Differentiation: A Dynamic Model for American Cities, 1960-1970." *Demography* 20:163-76.
- Silber, Jacques G. 1989. "On the Measurement of Employment Segregation." *Economics Letters* 30:237-43.
- Silén, Birgitta. 1988. "Women and Power." *Scandinavian Review* 76:91-101.
- Sobel, Michael E., Michael Hout, and Otis D. Duncan. 1986. "Saving the Bath Water. An Invited Comment on Krauze and Slomczynski's 'Matrix Representation of Structural and Circulation Mobility.'" *Sociological Methods and Research* 14:271-84.
- Sokoloff, Natalie J. 1987. "The Increase of Black and White Women in the Professions: A Contradictory Process." Pp. 53-72 in *Ingredients for Women's Employment Policy*, edited by Christine Bose and Glenna Spitze. Albany: State University of New York Press.
- Stafford, M. Therese, and Mark A. Fossett. 1989. "Occupational Sex Inequality in the Nonmetropolitan South, 1960-1980." *Rural Sociology* 54:169-94.
- Statistics Bureau of Japan. 1984. *1980 Population Census of Japan*, vol. 4, pt. 1. Tokyo: Statistics Bureau.
- Stier, Haya, and David B. Grusky. 1990. "An Overlapping Persistence Model of Career Mobility." *American Sociological Review* 55:736-56.

- Stolzenberg, Ross M., and Ronald J. D'Amico. 1977. "City Differences and Nondifferences in the Effect of Race and Sex on Occupational Distribution." *American Sociological Review* 42:937-50.
- Strober, Myra H. 1984. "Toward a General Theory of Occupational Sex Segregation: The Case of Public School Teaching." Pp. 144-56 in *Sex Segregation in the Workplace*, edited by Barbara F. Reskin. Washington, D.C.: National Academy.
- Szelényi, Szonja. 1988. *Social Mobility and Class Structure in Hungary and the United States*. Ph.D. dissertation. University of Wisconsin-Madison, Department of Sociology.
- Szelényi, Szonja, and Winifred R. Poster. 1991. "The Transformation of the Hungarian Class Structure, 1943-1980." Unpublished manuscript. Stanford University, Department of Sociology.
- Taeuber, Karl E., and Alma F. Taeuber. 1965. *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago: Aldine
- Tienda, Marta, and Vilma Ortiz. 1987. "Intraindustry Occupational Recomposition and Gender Inequality in Earnings." Pp. 23-51 in *Ingredients for Women's Employment Policy*, edited by Christine Bose and Glenna Spitze. Albany: State University of New York Press.
- Tienda, Marta, Shelley A. Smith, and Vilma Ortiz. 1987. "Industrial Restructuring, Gender Segregation, and Sex Differences in Earnings." *American Sociological Review* 52:195-210.
- Treiman, Donald J. 1977. *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
- U.S. Bureau of the Census. 1984. "Detailed Population Characteristics." *1980 Census of Population*, vol. 1. Washington, D.C.: Government Printing Office.
- Watts, Martin. 1992. "How Should Occupational Sex Segregation be Measured?" *Work, Employment, and Society* 6:475-87.
- White, Michael J. 1983. "The Measurement of Spatial Segregation." *American Journal of Sociology* 88:1008-18.
- Williams, Gregory. 1979. "The Changing U.S. Labor Force and Occupational Differentiation by Sex." *Demography* 16:73-87.
- Willms, Angelika. 1982. "Modernisierung durch Frauenarbeit? Zum Zusammenhang von wirtschaftlichem Strukturwandel und weiblicher Arbeitsmarktlage in Deutschland, 1882-1939." Pp. 37-71 in *Historische Arbeitsmarktforschung: Entstehung, Entwicklung und Probleme der Vermarktung von Arbeitskraft*, edited by Toni Pierenkemper and Richard Tilly. Göttingen: Vandenhoeck & Ruprecht.
- Winship, Christopher. 1977. "A Reevaluation of Indexes of Residential Segregation." *Social Forces* 55:1058-66.
- Xie, Yu. 1992. "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables." *American Sociological Review* 57:380-95.
- Yamaguchi, Kazuo. 1983. "The Structure of Intergenerational Occupational Mobility: Generality and Specificity in Resources, Channels, and Barriers." *American Journal of Sociology* 88:718-45.
- . 1987. "Models for Comparing Mobility Tables: Toward Parsimony and Substance." *American Sociological Review* 52:482-94.
- Zoloth, Barbara S. 1976. "Alternative Measures of School Segregation." *Land Economics* 52:278-98.



# Vote Turnout of Nineteenth Amendment Women: The Enduring Effect of Disenfranchisement<sup>1</sup>

Glenn Firebaugh  
*Pennsylvania State University*

Kevin Chen  
*Columbia University*

"Nineteenth Amendment women" are women in the United States who came of age during or just after the era when women could not vote. The roughly 4,000 such women included in the National Election Studies of 1952–88 provide an unusual opportunity for testing whether general historical conditions during childhood and adolescence have enduring effects. Despite common claims of cohort effects, some scholars remain skeptical because cohort effects are notoriously difficult to distinguish from age and period effects. Nineteenth Amendment women were in fact less likely to vote in the 1952–88 elections and this gender gap is unique to the amendment cohorts. These results provide strong evidence for the enduring effects of a cohort's historical conditioning.

Spurred by low voter turnout in the 1923 mayoral election in Chicago, C. E. Merriam and H. F. Gosnell conducted the "first sample survey recorded in the annals of American political science" (Bennett and Bennett 1987, p. 158; see also Merriam and Gosnell 1924). They discovered a gender gap in that election: Women in Chicago were much less likely to vote, a phenomenon they attributed to the recency of the extension of voting rights to women and to the still-powerful influence of attitudes against women's suffrage. The 1923 Chicago result paralleled the result

<sup>1</sup> This article is based on work supported by the National Science Foundation (SES-8909421). Support was also provided by the Population Research Institute, Pennsylvania State University, which has core support from the National Institute for Child Health and Human Development (grant 1-HD28262-01). We are indebted to Michael Berkman, Barrett Lee, R. S. Oropesa, and Susan Welch for suggestions. Address correspondence to Glenn Firebaugh, Department of Sociology, 212 Oswald Tower, Pennsylvania State University, University Park, Pennsylvania 16802. E-mail: firebaugh@pop.psu.edu.

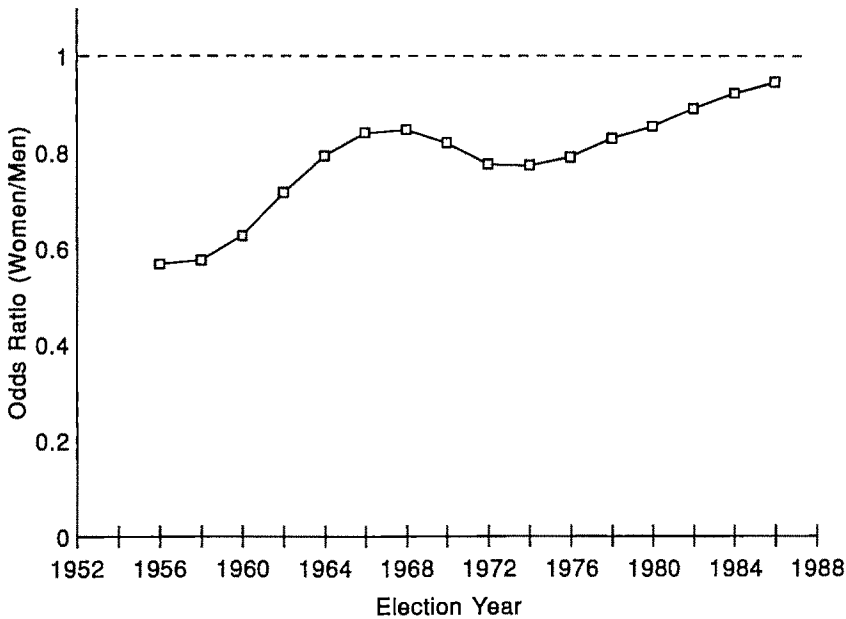


FIG. 1.—Odds of women voting versus odds of men voting in U.S. national elections (five-year moving averages).

for the whole of Illinois in the 1920 national election (Kleppner 1982, table 2; Illinois was the only state to report vote turnout separately for women and men).

In this study we ask whether women of the amendment era continued to exhibit substantially lower turnout rates than men 30–40 years after the passage of the Nineteenth Amendment. Available survey data since the middle of the century indicate a rather striking decline in sex differences in turnout for national elections (fig. 1). There are two distinct possibilities here. The decline might be across the board, with sex differences declining for all birth cohorts as the disenfranchisement era recedes further into history. Alternatively, disenfranchisement effects might apply only to those women who came of age before or during the amendment era (and not to their postamendment daughters and granddaughters), in which case the narrowing of the gender gap in vote turnout would result from the replacement of amendment-era women with post-amendment women.

The question of across-the-board convergence versus cohort-replacement effects bears on a long-standing debate over the importance of cohort effects for political behavior. At one end of a continuum are "generation theorists" who assert the primacy of political attitudes and behaviors

formed before adulthood (e.g., Abramson 1976; Inglehart 1971). At the other end are scholars who stress that attitudes are malleable throughout the life span (e.g., Lerner 1984), so that the imprint of history on adolescent cohorts is blurred as cohorts age and individuals change. Sociologists as well as political scientists have a stake in the issue (Alwin and Krosnick 1991). Though often posited, cohort effects are hard to demonstrate, because an identification problem (see below) makes the separation of age, period, and cohort effects notably difficult. Women's suffrage in the United States provides a rare opportunity to isolate cohort effects by testing for the enduring effects (if any) of the historical conditioning of presuffrage cohorts. Thus the telling question is: Other things being equal, did women who came of age before and just after suffrage exhibit depressed vote turnout rates over the life course?

#### DISENFRANCHISEMENT'S SIGNIFICANCE FOR COHORT THEORY

A long and hallowed tradition in the social sciences asserts that historical conditions present when one "comes of age" have enduring effects on one's attitudes and behavior (e.g., Mannheim 1952; see Ortega y Gasset [1933] for a related tradition in philosophy). Since all members of a birth cohort presumably come of age at roughly the same time, the effects of historical conditions become "embedded" in birth cohorts. If the effects persist throughout the cohort's life span, we say that there is a cohort effect. As Hobcraft, Menken, and Preston (1982, p. 10) put it, "Cohort effects occur whenever the past history of individuals exerts an influence on their current behavior in a way that is not fully captured by an age variable."

It is important to distinguish history-based from size-based cohort effects. "Size-based" refers to effects that arise because some cohorts are unusually large or small: "A cohort carries its fortunes, good or bad, depending on its size, throughout its life cycle" (Easterlin 1978, p. 404). Size differences in cohorts have been used to account for changes in, among other things, fertility, crime, earnings, and political alienation (e.g., Easterlin 1980; Welch 1979). History-based cohort effects, by contrast, result from the history—not the relative size—each cohort carries throughout its life cycle, because each cohort occupies a "unique location in the stream of history" (Ryder 1965, p. 844). That unique location produces cohort differences, because historical events "have a teaching impact . . . especially on those with less to unlearn" (i.e., the young; Stinchcombe 1984, p. 8). Put another way, historical events leave cohort-borne "residues." Demonstrating such residues empirically has proven to be tricky, however.

In contrast to life-cycle theories—which focus on the similarity be-

tween cohorts, as they more or less repeat the life-cycle patterns of their predecessors—cohort theories focus on cohort differences. As Ryder (1965, p. 844) explains, "Each new cohort makes fresh contact with the contemporary social heritage and carries the impress of the encounter through life. . . . Because it embodies a temporally specific version of the heritage, each cohort is differentiated from all others." Or as Lesthaeghe and Surkyn (1988, p. 40) put it, "Cohorts develop distinctive meaning-giving universes early in life and seem to maintain them throughout adulthood." Societal change arises from this cohort differentiation, as older cohorts are continually replaced by younger cohorts with different beliefs, values, and behavior (Firebaugh 1989). Some scholars of note believe that such "cohort replacement" accounts for a good deal of the social change observed in Western societies (e.g., Ryder 1965), so cohort effects occupy a prominent place in the literature on social change.<sup>2</sup>

Yet the question remains: Are cohort values and behaviors in fact shaped by the common youthful experiences of a cohort? There is no consensus on how prevalent and important cohort effects are, and some even question whether cohort effects exist at all. Among the most forceful defenders of cohort effects are students of recent European history, such as Inglehart (1971, 1985) and Lesthaeghe and Surkyn (1988), who attribute most of the recent social change in Europe to cohort effects (see also Weil 1987). Similar (though less sweeping) claims have been made for particular types of social and political change in the United States (e.g., Firebaugh and Davis 1988; Norpoth 1987). Panel studies of special populations (Bennington College students in the 1930s and 1940s [Alwin, Cohen, and Newcomb 1991] and 1960s activists [Jennings 1987]) also point to the possibility of cohort effects.

Others read the evidence differently. Schuman and Rieger (1992, p. 315) conclude that the "existence of cohort effects . . . has received little support from systematic research examining lasting associations between political events occurring during adolescence and early adulthood and present political and social attitudes and behavior." Though cohort effects have been found for "small activist and elite groups," findings "have been largely negative for the wider population." Duncan and Stenbeck (1988) are even more skeptical, arguing that a perceived cohort effect might well be an "artifact of a defective design" of the research (p. 15).

<sup>2</sup> A huge literature on political socialization (e.g., Sears 1990) is closely tied to cohort effects theory. Cohort effects also occupy a prominent place in the methods literature in developmental psychology, where they are generally viewed as a nuisance (Rosow 1978; Kosloski 1986).

The controversy persists because of an estimation problem. At any point in time, age is determined by date of birth; so in the absence of further information (such as direct measures of aging or life-cycle effects) one does not know whether observed cohort differences reflect the effect of aging or the effect of cohort membership (or both). The problem is not solved by adding data for other time periods, since that introduces the possibility of period effects, and cohort equals period minus age.

There are three basic ways to try to circumvent this estimation problem. The first is to impose *a priori* restrictions on the model. In the simplest form of this strategy, age effects (or period effects, or cohort effects) are assumed to be zero. Though more sophisticated versions of this strategy have appeared in recent years (see Sasaki and Suzuki [1987] for an example and for key references), the underlying logic is the same: Specified parameters are constrained in some way (to be zero, to be equivalent, etc.). Among others, Glenn (1976, 1989), Duncan and Stenbeck (1988), and Wilmoth (1990) are skeptical of this approach. Wilmoth (1990, p. 317) says, "I have thus demonstrated that the difficulties of modeling posed by perfect collinearity among the explanatory variables can have no purely statistical solution. Furthermore, the problem is not limited to the linear term but affects constant, quadratic, and all higher-order terms as well."

In the second approach—the one preferred by Converse (1976) and by Glenn (1989)—"side information" is used to judge the relative plausibility of the various combinations of age, period, and cohort effects that could mathematically account for an observed empirical pattern (see Firebaugh and Harley [1991] for an example). Nonlinear patterns can be especially telling. Suppose, for example, that one cohort differs sharply from those before and after and that the difference persists as the cohort ages. Strictly speaking, that pattern could arise from a particular combination of period and age effects (Wilmoth 1990), but the more plausible interpretation is that it reflects a true cohort effect. Though less apparent, the second strategy—like the first—involves assumptions; that is, certain combinations of age, period, and cohort effects are assumed to be more likely than others.

The third approach is to try to measure cohort effects directly. Though routinely applied in the case of size-based cohort effects, the direct-measurement strategy is rare in studies of history-based cohort effects (no doubt because history is hard to quantify).

Absent direct measures, the case for cohort effects often relies heavily on empirical evidence from two other sources. The first is the literature on the aging-stability thesis, the thesis that "with increasing age people become less likely to change" (Carlsson and Karlsson 1970, p. 710). On balance the evidence tends to support the thesis (Glenn 1980)—for some

attitudes, at least (Alwin and Krosnick 1991).<sup>3</sup> The second is the literature on cohort recall, which suggests that people disproportionately recall national events occurring when they were adolescents or young adults (Roberts 1986; Schuman and Scott 1989; Schuman and Rieger 1992). Nevertheless, as Roberts (1986, p. 84) cautions, it does not necessarily follow that remembered events are formative events, and Schuman and Rieger (1992) in fact find little relation between collective memory and attitudes.

In short, cohort effects remain somewhat controversial.<sup>4</sup> Despite their intuitive appeal, unambiguous examples are rare (Weil 1987), because it is hard to prove that a supposed cohort effect is not something else. Proving cohort effects may require, in the words of Schuman and Rieger (1992, p. 315), an "unusual conjunction of social conditions."

We believe the delayed enfranchisement of women in the United States provides the sort of "unusual conjunction" called for. In principle, the problem of isolating the effects of historical conditioning could be solved through a controlled experiment in which people of the same age were randomly assigned to different historical conditions during their formative years. Though such an experiment is not feasible, the important features of such an experiment are replicated in the "natural experiment" of women's disenfranchisement. The extension of voting rights to women in 1920 provides unusual leverage for determining the lasting effects of different historical conditions for men and women during their formative years (see the appendix).

## HYPOTHESES OF THE STUDY

### Hypothesis 1: Cohort-Specific Sex Differences in Vote Turnout

In the natural experiment provided by women's belated suffrage in the United States, sex is the assignment mechanism, "exposure to disenfranchisement" is the treatment (*T*), and voting in national elections is the

<sup>3</sup> Glenn (1980) provides a good summary of the theoretical arguments for the aging-stability thesis. First, there tends to be "dense spacing of significant life events in early adulthood" (p. 603) and wider spacing thereafter. Second, there may be an "incumbency" advantage, because new attitudes must be squared with old ones. Third, it is possible that physiological changes inherent in aging itself bring about increased rigidity.

<sup>4</sup> Some debates over cohort effects have arisen in a particular substantive context. Of these, probably the most famous is the Converse-Abramson debate. Consistent with the conclusion of *The American Voter* (Campbell et al. 1960), Converse (1976, 1979) interpreted age differences in political partisanship in the United States until the mid-1960s as reflecting age effects: As individuals age, they tend to become more partisan. Abramson (1976, 1979) disagreed, asserting the primacy of cohort effects in political partisanship. The controversy brought cohort analysis to center stage for a while in political science, and illustrated well the inherent difficulty of separating cohort effects from age and period effects.

criterion variable ( $Y$ ). The validity of a controlled experiment hinges largely on (1) random assignment—whether the assignment mechanism is truly random—and (2) uncontaminated treatment effect—whether the treatment the groups receive is equivalent, aside from  $T$  itself. With regard to condition 1, the sex of a baby is random. Condition 2, uncontaminated treatment effect, is more problematic. Because the different treatment of males and females in the pre-1920 United States was not confined to voting rights, observed differences in the later voting propensity of these two groups might reflect more than the effect of disenfranchisement itself. To alleviate this potential contamination problem, in addition to using control variables we use a treatment effect design keyed to specific categories of Nineteenth Amendment women (hypothesis 1) and contrast our results for voting with results for political variables other than voting (hypothesis 2).

Our treatment effect design is based on these categories of Nineteenth Amendment women: those born before 1896, those born from 1896 through 1905, and those born from 1906 through 1915. In contrast to women born after 1915—most of whom are not old enough to remember the enactment of the Nineteenth Amendment—women in our sample born before 1896 experienced delayed enfranchisement,<sup>5</sup> and women born 1896–1915 were socialized during an era when women could not vote. If disenfranchisement has enduring causal effects, its imprint should be greatest among the delayed enfranchisement cohorts and decline monotonically through the transition cohorts of 1896–1915 women who, though never actually denied the right to vote on the basis of their sex, were nevertheless socialized under such a regime. After passage of the Nineteenth Amendment, historical conditions no longer “taught” young women the impropriety of voting. So if historical conditioning is the telling factor, postamendment cohorts should exhibit no sex differences in vote turnout. Hence hypothesis 1:

**HYPOTHESIS 1.**—(a) *Women who were denied the right to vote on the basis of their sex or who were socialized when adult women could not vote will exhibit lower voting rates than men born during the same period, with the magnitude of the sex difference being greatest among those born before 1896, next greatest among those born 1896–1905, and less (but*

<sup>5</sup> The 1896 cutoff is based on presidential elections. With regard to electing a president, the 1896 cohort is the first one in which most men and women entered the electorate at the same age. Men born in 1895 could vote in the 1916 presidential election, but in most states women born in 1895 could not. So in most states men and women in the 1895 cohort did not enter the electorate together. However, men and women born in 1896 did enter the presidential electorate together. Those born in 1896, whether male or female, were first eligible to vote for a president in 1920 when they were 24 years old.

TABLE 1  
COHORT CATEGORIES AND SUMMARY OF PREDICTIONS

BIRTH-COHORT CATEGORY	N*		PREDICTIONS FROM HYPOTHESIS 1	
	Women	Men	Gender Gap†	Size of Difference
Before 1896 . . . . .	881	719	Yes	Largest
1896-1905 . . . . .	1,209	973	Yes	Second largest
1906-15 . . . . .	1,939	1,637	Yes	Third largest
1916-25 . . . . .	2,402	2,003	?	Smallest (or zero)
1926-35 . . . . .	2,089	1,737	No	
1936-45 . . . . .	1,807	1,507	No	
1946-55 . . . . .	1,831	1,506	No	
After 1955 . . . . .	937	796	No	
Total . . . . .	13,095	10,878		

\* The logit coefficients reported in table 2 are based on this sample

† Vote turnout for women is *lower* than vote turnout for men, other things being equal

*nonzero) among those born 1906-15. (b) There will be no sex difference in vote turnout among those socialized after the Nineteenth Amendment was enacted.*

This is a strong hypothesis, because we predict the relative magnitudes of the nonzero coefficients. Table 1 summarizes our predictions. As implied in the table, there are eight coefficients of consequence, one for each of the eight cohort categories. We expect negative coefficients in three instances (women's voting rates lower than men's for these cohort categories), nonsignificant coefficients in four instances (no difference in women's and men's rates), and in one instance—the "bridge" cohort category, consisting of those born 1916-25, just before and after the Nineteenth Amendment—we predict only that, if there are sex differences, they will be smaller than those in the previous cohort category.

It is important to stress that hypothesis 1 attempts to demonstrate the presence of cohort effects by predicting a cohort pattern keyed to the passage of the Nineteenth Amendment. We predict that the greater the "exposure" to disenfranchisement, the greater its cohort residue. Thus disenfranchisement should tend to have the greatest lasting effect on women actually denied the right to vote, with a lesser effect on women who were adolescents during the disenfranchisement era, followed by those who were children during that era.

#### Hypothesis 2: Differences Limited to Voting

Even if we find a pattern of sex differences keyed to the passage of the Nineteenth Amendment, per hypothesis 1, skeptics might nonetheless



aver that the amendment itself simply reflects more general cultural changes in the roles of women during the amendment era and that those general changes are the real causes. In other words, there is a history-based cohort effect, but the effect applies to a wide range of political attitudes and behaviors other than voting because it is rooted in broad cultural changes of which the extension of voting rights to women is but one manifestation.

In light of that possibility we also estimate sex differences, by cohort categories, for political variables other than voting. We hypothesize that the pattern of cohort-specific differences outlined in table 1 holds only for voting.

**HYPOTHESIS 2.**—*The pattern of sex differences predicted in hypothesis 1 does not hold for political variables other than voting.*

The National Election Studies (NES 1990) data set we use contains a wide variety of measures of political behaviors and attitudes. We chose three classes of items that we believe are pertinent here: political involvement other than voting—following the campaign on television or donating money; trust in government—believing that the government can be trusted to do what is right and not waste tax money; and feelings of personal efficacy with respect to politics—feeling that public officials care what “people like me” think and that “people like me” have a say in government. Political efficacy is the most likely exception to the hypothesis that the amendment/postamendment distinction bears only on sex differences in voting, since the failure to vote in itself could reduce one’s feelings of political efficacy.

## DATA, MODEL, AND MEASUREMENT OF VARIABLES

### Data

We use biennial survey data from the 1952–88 NES (1990) carried out by the University of Michigan. Data were obtained from face-to-face interviews of respondents selected using a national full-probability sample of all voting-age citizens living in households in the continental United States, except for those on military bases.

Eighteen surveys are used (the 1954 survey was omitted because age data were absent). Our analysis is restricted to whites, since barriers to nonwhite voting persisted after 1920. If there is a disenfranchisement effect along the lines described earlier, it would apply to whites but not necessarily to nonwhites. For the models we estimate, complete data are available for roughly 24,000 whites (table 1). Our analysis is based on those 24,000 cases.

## Principal Dependent Variable: Self-Reported Voting

The NES is designed to correspond with nationwide elections in the United States. In the postelection surveys, respondents are asked if they voted—the principal dependent variable in this analysis.

Self-reported voting tends to be inflated. Voting rates calculated from the NES survey responses consistently exceed official voting rates, with overreporting rates typically in the 10%–15% range (Katosh and Traugott 1981; Sigelman 1982). The NES responses and official data nevertheless yield the same aggregate patterns for 1952–88 presidential elections: Turnout peaked in 1960, declined steadily until 1980, increased slightly in 1984, and declined again in 1988 (Chen 1992, fig. 1.1).

For the purposes of this study, misreporting of voting is not critical. By comparing results for validated and self-reported voting, Katosh and Traugott (1981) and Sigelman (1982) conclude that using self-reported data introduces no serious bias in studies of the determinants of voting (but see Silver, Anderson, and Abramson [1986] for qualifications). To anticipate a bit, we find that Nineteenth Amendment women do report lower voting behavior. That could mean either that they were in fact less inclined to vote or that they were more reluctant to admit they did (or both). Either way, we conclude that historical conditioning in the early 1900s has effects that persist into the second half of the century. Moreover, if disenfranchisement made amendment women more reluctant than men to admit that they did *not* vote, the observed gender gap *understates* the actual gap, implying a conservative test of our thesis.<sup>6</sup>

## Dependent Variables for Hypothesis 2

*Political involvement.*—We use two NES items to measure political involvement other than voting. The first is, “Did you watch any programs about the campaign on television?” (WatchTV; yes = 1).<sup>7</sup> The second is, “Did you give any money or buy any tickets or anything to help the campaign for one of the parties or candidates?” (Donate; yes = 1).

<sup>6</sup> We thank an *AJS* reviewer for pointing out this idea to us.

<sup>7</sup> The wording varies somewhat from survey to survey (the effect of wording change on the mean is captured by the election-year dummies). The fact that televisions were not ubiquitous at the time of the earliest elections here does not undermine our use of television watching to measure sex differences in political involvement, since there is no reason to believe that women were substantially less (or more) likely than men to have had televisions in their homes in the 1950s.

*Trust in government.*—We use these NES items to measure trust in government: “How much of the time do you think you can trust the government in Washington to do what is right—just about always, most of the time, or only some of the time?” (DoRight; just about always or most of the time = 1). “Do you think people in the government waste a lot of money we pay in taxes, waste some of it, or don’t waste very much of it?” (NoWaste; waste some or don’t waste very much = 1).

*Personal political efficacy.*—We use these NES items to measure feelings of efficacy with respect to the polity: “People like me don’t have any say about what the government does” (HaveSay; disagree = 1). “I don’t think public officials care much what people like me think” (GovCare; disagree = 1).

### Modeling the Hypothesized Treatment Effect

As explained above, the different voting treatment of men and women born in the 19th and early 20th centuries provides rare leverage for distinguishing cohort effects from period and age effects. As hypothesis 1 and the appendix make clear, the key is to focus on *cohort-specific* gender gaps.

The estimation of cohort-specific gender gaps requires two sets of dummy variables: one set of seven dummies for the eight cohort categories and another set of eight dummies for sex differences (cohort category times sex, where sex is coded “women = 1”). That is, we estimate this model:

$$\text{logit}(\text{vote}) = \alpha + C\gamma + C_w\delta + Z\beta + \epsilon, \quad (1)$$

where  $C$  is a  $1 \times 7$  vector coded [0000000] for those in cohort category 1 (born before 1896), [1000000] for those in category 2, [0100000] for those in category 3, . . . , [0000001] for those in category 8;  $C_w$  is a  $1 \times 8$  vector coded [00000000] for men, [10000000] for women in cohort category 1, [01000000] for women in category 2, . . . , [00000001] for women in category 8;  $Z$  is a vector of covariates; and  $\gamma$ ,  $\delta$ , and  $\beta$  are column vectors of parameters. The important point here is that the  $\delta$ s capture differences between men and women in the same cohort category.<sup>8</sup>

<sup>8</sup> This point is readily seen by writing out eq. (1) with the covariates set at zero. For men in cohort category 1 the predicted  $\text{logit}(\text{vote})$  is  $\alpha$ ; for men in category 2 the prediction is  $\alpha + \gamma_1$ ; for men in category 3,  $\alpha + \gamma_2$ ; and so on. So the  $\gamma$ s reflect cohort differences in voting behavior *for men*—the additive cohort effect here. For women in cohort category 1 the predicted  $\text{logit}(\text{vote})$  is  $\alpha + \delta_1$ ; for women in category 2 the prediction is  $\alpha + \gamma_1 + \delta_2$ ; for women in category 3,  $\alpha + \gamma_2 + \delta_3$ ; etc. So the  $\delta$ s capture differences in vote turnout (covariates controlled) for men and women, cohort by cohort:  $\delta_1$  is the difference for men and women born before 1896,  $\delta_2$  is the difference for men and women born 1896–1905, etc.

## Covariates

*Period.*—Turnout tends to be higher for presidential elections. Turnout also no doubt is affected by election-specific factors (e.g., Kleppner 1982), such as attractiveness of the candidates, perceived differences between the candidates, closeness of the election, campaign spending, interest generated by the campaigns, the state of the economy, and the weather on election day. Dummy variables are included to capture any such period or “election” effects.

*Age.*—It is well established that age affects voting, with younger adults less likely to vote than middle-aged adults (e.g., Campbell et al. 1960, chap. 17; Hout and Knoke 1975). Voting rates level off at middle age and decline after age 65 or 70 (Chen 1992, fig. 1.3). Age and age squared are included to capture this quadratic effect.

*Education.*—Formal education has a positive effect on voting (e.g., Wolfinger and Rosenstone 1980). Controlling for education is critical here, lest the effect of disenfranchisement be confounded with the effect of historical differences between men and women in the amount of formal education acquired. The NES education variable has six categories, ranging from eight or fewer years of schooling to college degree.

*Family income.*—Some studies find that income has an effect on voting participation, independent of the effect of education (e.g., Presser and Traugott 1992). Family income in the NES is a relative measure. That is, respondents are placed in one of five categories based on their family income relative to the income of other families at that time.

*Region.*—Even after the effects of race and socioeconomic status are controlled for, southerners are less likely to vote in national elections. Rusk and Stucker (1978) attribute this effect to the southern system of election laws enacted in the late 19th century, which effectively reduced the voting of whites (especially poor whites) as well as of blacks (but see Key 1949, chap. 25). Whatever its underlying cause, this regional effect has persisted to the present (Chen 1992, table 8.2), so we include a dummy variable (South = 1) to control for it.<sup>9</sup>

*Strength of party partisanship.*—Strong party partisans are more likely to vote (Campbell et al. 1960, chap. 5). Strength of party partisanship is measured on a four-point scale, from “independent or apolitical = 1” to “strongly partisan = 4.”

*Marital status.*—Because life expectancy has increased over the last century, widowhood tends to come later in life. Hence if we compare

<sup>9</sup> Persons were classified as southern if they lived in Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia, or Washington, D.C.

two women at, say, age 65—one born before 1916, the other born later—the woman born before 1916 is more likely to be a widow. Since the unmarried are less likely to vote (Petrocik 1991), lower turnout for women born before 1916 might reflect marital effects of the present rather than cohort (disenfranchisement) effects rooted in the past. To eliminate that possibility, we include marital status (married = 1) as a covariate.

## FINDINGS

### Results for Hypothesis 1

Table 2 reports results for our test of hypothesis 1. We use logistic regression, since the dependent variable is a dichotomy. Estimates are based on the biennial nationwide elections in the United States, 1952–88 (except 1954, as noted earlier).

Results are consistent with the hypothesis. Controlling for the effects of period (election year), age, education, marital status, and so forth, Nineteenth Amendment women are in fact less likely to vote than men born during the same period, with the sex differences being greatest for the most affected cohort category and declining monotonically thereafter. Thus, cohorts with the most enduring exposure to disenfranchisement exhibit the greatest sex differences. It is important to note that post-amendment men and women vote at the same rate.

The magnitude and pattern of the disenfranchisement effect is noteworthy. Among those born before 1896, a woman's odds of voting are less than half (.47) that of a man's odds, other things being equal.<sup>10</sup> The disenfranchisement effect is only slightly less for women born from 1896 through 1905. Women in this cohort category were never themselves denied the right to vote on the basis of sex, yet they came of age during a period when women could not vote. The disenfranchisement effect is notably smaller for those in the next cohort category (born 1906–15). The size of the reduction—from  $-0.68$  to  $-0.36$  in the logits (table 2)—squares with the assumption of some cohort theories that adolescent years are especially important in the formation of political attitudes and behavior. Most women born from 1906 through 1915 were old enough to remember when women could not vote, yet the prohibition was lifted before they reached their most formative years (according to Mannheim, at "the age of 17, sometimes a little earlier and sometimes a little later" [1952, p. 300]).

<sup>10</sup> It is interesting to compare these odds with those obtained from the 1920 Illinois data cited earlier. In 1920, being a woman reduced the odds of voting by two-thirds (calculated from Kleppner [1982, table 2]). To be sure, this result is based on a single state, and there are no control variables. The result nonetheless suggests that historical conditioning effects might tend to fade somewhat over time.

TABLE 2

COHORT-SPECIFIC SEX DIFFERENCES IN VOTING IN NATIONAL ELECTIONS, 1952-88:  
ESTIMATES FOR LOGISTIC REGRESSION MODEL WITH COVARIATES

	Logit	Odds Ratio	P-value
<b>Sex differences by birth-cohort category:</b>			
Before 1896 .....	-.76	.47	<.0001
1896-1905 .....	-.68	.51	<.0001
1906-15 .....	-.36	.70	.0001
1916-25 .....	-.25	.78	.002
1926-35 .....	-.04	.96	.61
1936-45 .....	.06	1.06	.47
1946-55 .....	-.03	.97	.68
After 1955 .....	-.06	.95	.61
Combined dummies* .....	.	...	<.0001
<b>Covariates:</b>			
Age† .....	.07	1.08	<.0001
Age‡ .....	-.001	.999	<.0001
Education .....	.50	1.66	<.0001
Family income .....	.22	1.24	<.0001
Region .....	-.66	.51	<.0001
Party partisanship .....	.42	1.52	<.0001
Marital status .....	.21	1.23	<.0001
Birth-cohort-category dummies§ ...	...	...	.02
<b>Election-year dummies:§</b>			
1956 .....	-.20	.82	.06
1958 .....	-.99	.37	<.0001
1960 .....	.30	1.35	.02
1962 .....	-1.01	.36	<.0001
1964 .....	-.06	.94	.61
1966 .....	-.96	.38	<.0001
1968 .....	-.10	.90	.46
1970 .....	-.97	.38	<.0001
1972 .....	-.15	.86	.27
1974 .....	-.94	.39	<.0001
1976 .....	-.24	.79	.13
1978 .....	-1.07	.34	<.0001
1980 .....	-.28	.76	.12
1982 .....	-1.04	.35	<.0001
1984 .....	-.18	.84	.35
1986 .....	-1.36	.26	<.0001
1988 .....	-.45	.63	.03
Combined dummies¶ .....	.	...	<.0001

NOTE —Coefficients estimated using listwise deletion of missing cases;  $N = 23,973$  whites. Observed voting = 69.19%, number correctly predicted by model = 74.82%, proportional reduction in error = 183.

\* Partial- $r = .052$  for sex differences collectively. Significance test is based on Wald statistic.

† Age is coded as actual age minus 17.

‡ For the cohort dummies collectively, partial- $r$  is trivial ( $r = .010$ ) and, despite the large sample, fails to attain significance at the  $P < .01$  level. Thus we do not report logits for individual cohort dummies.

§ The omitted (reference) category is the 1952 election.

¶ Partial- $r = .160$

Despite the large sample, among those born after 1925 there are no statistically significant differences between men's and women's electoral participation. That disappearance of sex differences in voting among postamendment cohorts buttresses the view that sex differences among earlier cohorts in fact reflect the imprinting effect of women's disenfranchisement. Women socialized during the period when women could not vote in national elections vote at lower rates than men, whereas women socialized later exhibit the same voting rate as men. The sole exception is the 1916–25 cohort category—the one straddling the Nineteenth Amendment—where women socialized just after the disenfranchisement era are somewhat less likely to vote than men.

Results for the covariates are in line with expectations. The age effect is quadratic (age and age squared are both statistically significant, despite their high correlation:  $r = .965$ ; see table A1 in the appendix), with the middle-aged and old more likely to vote than the young and the very old. Education, income, political partisanship, and marriage have independent positive effects on self-reported voting. Southerners are less likely to report voting in nationwide elections, other things being equal.

With one exception, all variables in table 2 display effects that are significant at  $P \leq .0001$ . The sole exception is the additive cohort effect, with a  $P$ -value of only .02 (the partial- $r$  is only .010). So cohort differences in *men's* electoral participation are unimportant and, to save space in table 2, we do not report logits for the cohort dummies. Table 2 does report logit coefficients for the 17 election-year dummies (period effects). Note the striking pattern of alternating significant and nonsignificant coefficients, reflecting the reduced turnout for nonpresidential elections in the United States. Nevertheless, the big story here is neither period effects nor (additive) cohort differences but the importance of voting differences between men and women who lived during the disenfranchisement era.

*Robustness of the disenfranchisement effect.*—Some readers might wonder how much our results would change if we changed the cohort categories. The most telling way to change the categories is to alter the starting point. We chose 1896 as the starting point for our categorization in table 2, because 1896 is the first cohort where men and women in all states entered the electorate together (n.5, above). Though a few states allowed women to vote in national elections before 1920, one could argue that it is national, not state, norms that are important here, and until 1920 all women were subject to the national “disenfranchisement norm,” even if some could in fact vote. In any case, to check the robustness of our results we reestimated table 2 using 1894, then 1892, and then 1890 as the starting point for categorizing cohorts.

Regardless of how we categorize the cohorts, we find that amendment

women were less likely than their male contemporaries to vote in the 1952–88 national elections (results available on request), with the gender gap being largest among the cohorts most affected by disenfranchisement and disappearing among the postamendment cohorts. Upon close inspection our results serendipitously support Mannheim's "formative-years" thesis, since the sharpest decline in sex differences appears to be linked to adolescence in 1920. Specifically, sex differences decline more sharply just after, not before,<sup>11</sup> the cohort categories containing women who were 12–18 years old in 1920. So adolescence appears to be the key life stage—a result that is more in line with Mannheim's coming-of-age version of cohort theory than with Ryder's (1965) version, which stresses the effects of socialization from birth onward.

### Results for Hypothesis 2

Table 3 reports results for our test of hypothesis 2. In the case of voting, there is a gender gap that is unique to the amendment cohorts. The question is: Are other political variables also characterized by such a pattern?

The answer is no. The pattern of a large but declining gender gap among amendment cohorts, followed by no gender gap among post-amendment cohorts, fails to hold for political variables other than voting (see table 3). Of the six items we compare to voting, only the pattern of sex differences for the HaveSay question even remotely resembles the pattern of sex differences in voting behavior. That is, Nineteenth Amendment women are less likely than their male contemporaries to feel that they have any say about what the government does—which is reasonable, since they are also less likely to vote—yet postamendment women do not differ from their male contemporaries in their responses to this question—again reasonable, since men and women in these cohorts are equally likely to vote. Even for the HaveSay question, however, the mimicking of voting patterns is not exact (compare cols. 1 and 6 of table 3), since for HaveSay there is no monotonic decline in sex differences among the amendment cohorts.

For the other political variables, we notably fail to replicate the amendment/postamendment cohort pattern found for voting. Regardless of when the persons were born, there is no statistically significant difference in the responses of women and men as to whether or not the government

<sup>11</sup> The cohort category just before the adolescent cohort category contains those who were actually denied voting rights because of their sex. Hence one might expect to see the greatest decline in the disenfranchisement effect between that category and its successor, the category of adolescents. Yet such is not the case.



TABLE 3  
COHORT-SPECIFIC SEX DIFFERENCES IN POLITICAL BEHAVIORS AND ATTITUDES: ESTIMATED ODDS RATIOS  
FOR LOGISTIC REGRESSION MODEL WITH COVARIATES

BIRTH-COHORT CATEGORY	POLITICAL INVOLVEMENT		TRUST IN GOVERNMENT		PERSONAL POLITICAL EFFICACY		
	Vote	WatchTV <sup>a</sup>	Donate <sup>a</sup>	DoRight <sup>a</sup>	NoWaste <sup>a</sup>	HaveSay <sup>a</sup>	GovCare <sup>a</sup>
Before 1896 .....	.47**	.87	.84	.99	1.22	.76*	.80
1896-1905 .....	.51**	1.35*	.75	1.00	1.11	.68**	.93
1906-15 .....	.70**	1.21	.74*	.96	1.13	.86	.95
1916-25 .....	.78*	.92	.87	.99	1.09	.88	.98
1926-35 .....	.96	1.08	.77*	.86	1.15	.95	1.13
1936-45 .....	1.06	.81	1.13	1.10	1.59**	1.00	1.33**
1946-55 .....	.97	.93	.66**	.89	1.35**	1.06	1.25*
After 1955 .....	.95	.79*	.51*	.82	1.18	.98	1.05
P-value <sup>b</sup> .....	<.0001	.027	<.0001	.19	<.0001	.001	.0003
Partial- <sup>a</sup> .....	.052	.009	.037	.000	.040	.019	.021
N .....	23,973	16,807	21,248	18,370	16,558	20,355	20,732

NOTE.—Odds ratios (women's odds/men's odds), controlling for election year, age, age<sup>2</sup>, cohort category, education, family income, region, party partisanship, and marital status (as in table 2). Data are from the 1952-88 National Election Studies (NES 1990). Results for covariates are suppressed to save space. To facilitate comparison, the first column repeats results for voting odds ratios shown in table 2.

\*  $P < .05$ .

\*\*  $P < .001$ .

<sup>a</sup> For definitions of dependent variables, see the text section entitled "Dependent Variables for Hypothesis 2."

<sup>b</sup> For cohort-specific sex differences collectively (significance test is for the overall Wald statistic).

can be trusted to do what is right “just about always” or “most of the time” (DoRight). Similarly, there is no difference of consequence between women and men with regard to following campaigns on television (WatchTV; sample differences fail to reach statistical significance at  $P < .01$  [ $P = .027$ ], despite a sample size of nearly 17,000). There are sex differences in political giving (Donate), but this gender gap apparently has nothing to do with disenfranchisement, since women tend to be less likely than men to donate regardless of when they were born. The amendment/postamendment cohort distinction is likewise immaterial to sex differences in response to the question about wasting tax money (NoWaste). Whether or not they were socialized during the disenfranchisement period, women are less likely than men to say that people in government waste a lot of money.

Our failure to replicate the voting results for nonvoting variables adds credibility to the view that the reduced voting of amendment women from 1952–88 is rooted in their exposure to disenfranchisement itself. There is no hint in these data that—aside from voting—amendment women tend to be “less political” than their male contemporaries in ways that also distinguish them from postamendment women.<sup>12</sup> We conclude, in short, that the reduced electoral participation of Nineteenth Amendment women stems from the conditioning effect of disenfranchisement itself.

#### Implications for Trends in the Gender Gap in Voting

Historically women have had lower voting rates than men (Campbell et al. 1960, chap. 17; Welch 1977), but the gender gap has evaporated in recent years (see fig. 1). Our findings suggest a straightforward explanation for this convergence: cohort replacement. If sex differences in turnout are restricted to amendment cohorts, then the overall gender gap in turnout will disappear as these cohorts disappear.

We addressed the replacement issue more directly by reestimating the vote-turnout model separately for the first six, middle six, and last six elections from 1952–88. The telling question here is whether cohort-specific sex differences declined noticeably over time. If cohort-specific sex differences did not decline over time, then the decline in the overall gender gap must be due to cohort replacement, that is, to an increase in the proportion of the population in the postamendment cohorts (where there are no sex differences in voting rates).

<sup>12</sup> In the case of political donations, e.g., amendment women tend to be “less political” (contribute less) than their male contemporaries, but this holds true for post-amendment women as well. Thus, in the case of political donations, the gender gap is not restricted to amendment women.

TABLE 4

COHORT-SPECIFIC SEX DIFFERENCES IN VOTING. ESTIMATES FOR EARLIER,  
MIDDLE, AND LATER NATIONAL ELECTIONS

BIRTH- COHORT CATEGORY	1952-64		1966-76		1978-88	
	Odds Ratio <sup>a</sup>	N	Odds Ratio <sup>a</sup>	N	Odds Ratio <sup>a</sup>	N
Before 1896 .....	.43**	1,126	.59*	435	.37	39
1896-1905 .....	.46**	1,083	.51**	808	.81	291
1906-15 .....	.63**	1,533	.74*	1,304	.73	739
1916-25 .....	.60**	1,804	.90	1,540	.99	1,061
1926-35 .....	1.07	1,272	.83	1,442	1.01	1,112
1936-45 .....	1.33	351	1.07	1,579	.97	1,384
1946-55 .....	...	0	.92	1,176	1.00	2,161
After 1955 .....	...	0	1.10	83	.93	1,650
Total N .....		7,169		8,367		8,437
P-value <sup>b</sup> ..	<.0001		<.001		.79	
Partial- $r^b$ .....	.088		.032		<.00005	

NOTE.—Results are from the 19 national elections in the United States, 1952-88, except 1954 (data unavailable).

\*  $P < .05$ .

\*\*  $P < .001$ .

<sup>a</sup> Voting odds ratios (odds of women voting/odds of men voting), controlling for election year, age, age<sup>2</sup>, cohort category, education, family income, region, party partisanship, and marital status (as in table 2). Results for covariates are suppressed to save space.

<sup>b</sup> For cohort-specific sex differences collectively.

Our results (table 4) support the cohort-replacement account. The vanishing gender gap is due primarily to the diminution of amendment cohorts, not to the diminution of sex differences within those cohorts. Note first the striking decline in the importance of cohort-specific sex differences from 1952-88: partial- $r$  declines from .088 ( $P < .0001$ ) for the first six elections to .032 ( $P < .001$ ) for the middle six to <.00005 ( $P = .79$ ) for the last six. This result suggests either that sex differences have evaporated for the amendment cohorts or that sex differences in the amendment cohorts no longer matter much because those cohorts have come to make up such a small proportion of the population (or both). The former implies odds ratios of 1.0 for the amendment cohorts in the later elections, yet the NES data indicate otherwise. For the three amendment-cohort categories (born before 1896, born 1896-1905, born 1906-15), but not for the bridge category (born 1916-25), the odds ratios for the sample remain consistently below 1.0.

Thus the decline in the overall gender gap must result primarily from the changing cohort composition of the U.S. electorate, as the Nineteenth Amendment cohorts gradually die off and are replaced by cohorts not conditioned by disenfranchisement. The sample sizes in table 4 vividly

demonstrate the shrinking relative size of the amendment cohorts over the 1952–88 period. For the 1952–64 elections, the majority of the electorate—3,742 of 7,169 (52%)—had been socialized during the disenfranchisement period. For the 1966–76 elections, the amendment cohorts had shrunk to 30% of the total; by the 1978–88 elections, members of amendment cohorts constituted only about one-eighth (13%) of the NES sample. Since sex differences in voting rates are restricted to those persons socialized before and just after enactment of the Nineteenth Amendment, voting rates converge as those cohorts disappear.

### SUMMING UP

Because the right to vote was extended to women more than three decades before the elections studied here, our results indicate the staying power of norms internalized many years earlier. Historical conditions during formative years can have long-lasting effects.

Yet, because conditioning effects are embedded in cohorts, they do not last forever. The effects disappear when the cohorts disappear. With regard to vote turnout, disenfranchisement had enduring pernicious effects on Nineteenth Amendment women but not on their postamendment daughters and granddaughters. As a result, voting rates for men and women have converged.

### APPENDIX

#### The Delayed Enfranchisement of Women as an Unusual Opportunity to Demonstrate History-Based Cohort Effects

In this appendix we show more formally how the delayed enfranchisement of U.S. women provides an unusual conjunction of conditions that permits a plausible demonstration of cohort effects. In addition, we spell out how our methods here relate to conventional methods for circumventing the age-period-cohort (APC) identification problem.

Disenfranchisement effects imply cohort-specific sex differences that follow a predicted pattern. To relate the argument to the earlier literature on APC effects, consider separate equations for the vote turnout of men and women. Since nothing essential is lost if we ignore variables other than age, period, and cohort, we focus on the reduced (no covariates) model:<sup>13</sup>

$$v_{ijk}^M = \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}^M$$

<sup>13</sup> An *AJS* reviewer suggested this way to frame the issue. Our discussion of model (A1) borrows liberally from the reviewer's comments to us.

TABLE A1  
CORRELATIONS OF VARIABLES WITH MEANS AND SDs

	Vote	Age <sup>a</sup>	Age <sup>2</sup>	Female	Education	Family Income	South	Party Partisan	Married
Vote .....	. . .								
Age* ..	.14								
Age <sup>2</sup> .....	.09	.965							
Female .....	-.06	.01	.02						
Education .....	.18	-.31	-.31	-.03					
Family Income .....	.19	-.26	-.34	-.12	.41				
South .....	-.15	.01	.02	.00	-.06	-.10			
Party partisan .....	.20	.20	.18	.03	-.04	-.01	.01		
Married ..	.10	-.14	-.20	-.13	.00	.36	-.01	.01	. . .
Mean .....	.692	28.2	1,071	.546	2.96	2.99	.299	2.84	720
SD .....	.462	16.7	1,093	.498	1.29	1.12	.458	.98	.449

NOTE.—N = 23,973  
\* Actual age minus 17

and

(A1)

$$v_{ijk}^W = \alpha_i + \beta_j + \gamma_k + \delta_k + \epsilon_{ijk}^W$$

where  $v_{ijk}^M$  is the logit(proportion voting) for men age  $i$ , in period  $j$ , from cohort  $k$  ( $k = j - i$ ), and similarly for women ( $v_{ijk}^W$ );  $\alpha_i$ ,  $\beta_j$ , and  $\gamma_k$  are the *common* age, period, and cohort effects for men and women,  $\delta_k$  is the cohort difference, and the usual assumptions are made about the  $\epsilon$ s.

The variable of interest in this analysis is  $\delta_k$ , the cohort-specific sex difference term. There is no identification problem in estimating  $\delta_k$ , because

$$v_{ijk}^W - v_{ijk}^M = \delta_k + (\epsilon_{ijk}^W - \epsilon_{ijk}^M). \quad (A2)$$

Model (A2) avoids the identification problem by "differencing out" the common age, period, and cohort effects for men and women. The key assumption, then, is that  $\alpha_i$  and  $\beta_j$  are the same for men and women. (Except where men and women are analyzed separately, or modeled with interaction terms, that assumption is made implicitly anyway in cohort analyses of men and women.)

Our objective in this analysis is to estimate the  $\delta_k$ 's and show that they are related to disenfranchisement. There are two potential pitfalls. The first is that the  $\alpha_i$ 's and  $\beta_j$ 's might in fact differ for men and women. Age effects might be different for men and women, for example, if at older ages women are more likely than men to be widowed and if widowhood affects the probability of voting. We address this possible problem by including covariates such as marital status in our analysis.

The second potential problem is this: Perhaps the  $\delta_k$ 's truly are cohort effects but ones that reflect educational differences between men and women, general cultural change, or something else other than disenfranchisement. We address this problem by including covariates such as education, by showing that the pattern of the sex differences is keyed to the passage of the Nineteenth Amendment (see hypothesis 1), and by showing that the sex differences occur *only* in voting (hypothesis 2).

How do the methods we used relate to conventional approaches to the identification problem? We need to note first that we actually did two analyses of vote turnout, one based on the eight cohort categories of table 1, the other based on equation (A2), where each of the 104 birth years is treated as a separate cohort. Either way we reach the same conclusions about the  $\delta_k$ 's. We report the results for the broader cohort categories, because it would be unwieldy to report over 100 sex differences, some of which are based on trivial  $N$ 's.

In the case of equation (A2)—the "first-difference" (Liker, Augustyniak, and Duncan 1985) or "fixed effects" (England et al. 1988) solu-

tion—APC effects are differenced out, as already noted. In the results reported, where broader cohort categories are used, it is no longer strictly true that cohort equals period minus age. Using broad cohort categories is a common strategy in cohort analysis; in effect it replaces perfect collinearity with high collinearity. Because the cohort categories often are labeled (e.g., "Depression cohort" for those who came of age during the Great Depression), one might be tempted to classify this as a measurement strategy, since cohort is "measured" by grouping adjacent birth years whose members were exposed to a common momentous historical event during their formative years. In point of fact, though, the method handles the identity problem by placing a priori restrictions on some of the parameters: The effect of birth year, independent of age and period, is constrained to be equivalent for groups of adjacent birth years.

Except for the first and last cohort categories, we use 10-year categories (born 1896–1905, etc.). We reach the same conclusions when using 10-year categories beginning at different points, as noted in the text. Though this consistency is reassuring, there is no definitive way to test the identifying restrictions imposed by the broadening of cohort categories. For that reason we are less confident about our estimates of the  $\alpha_i$ 's,  $\beta_j$ 's, and  $\gamma_k$ 's than we are about our estimates of the  $\delta_k$ 's.

## REFERENCES

- Abramson, Paul. 1976. "Generational Change and the Decline of Party Identification in America: 1952–1974." *American Political Science Review* 70:469–78.
- . 1979. "Developing Party Identification: A Further Examination of Life-Cycle, Generational, and Period Effects." *American Journal of Political Science* 23:78–96.
- Alwin, Duane F., and Jon A. Krosnick. 1991. "Aging, Cohorts, and the Stability of Sociopolitical Orientations over the Life Span." *American Journal of Sociology* 97:169–95.
- Alwin, Duane F., Ronald L. Cohen, and Theodore M. Newcomb. 1991. *Political Attitudes over the Life Span: The Bennington Women after Fifty Years*. Madison: University of Wisconsin Press.
- Bennett, Stephen, and Linda Bennett. 1987. "Political Participation." *Annual Review of Political Science* 2:157–204.
- Campbell, Angus, Philip Converse, Warren Miller, and Donald Stokes. 1960. *The American Voter*. New York: Wiley.
- Carlsson, Gosta, and Katarina Karlsson. 1970. "Age, Cohorts, and the Generation of Generations." *American Sociological Review* 35:710–18.
- Chen, Kevin. 1992. *Political Alienation and Voting Turnout in the United States, 1960–1988*. San Francisco: Mellen Research University Press.
- Converse, Philip E. 1976. *The Dynamics of Party Support: Cohort-Analyzing Party Identification*. Beverly Hills, Calif.: Sage.
- . 1979. "Rejoinder to Abramson." *American Journal of Political Science* 23: 97–100.
- Duncan, Otis Dudley, and Magnus Stenbeck. 1988. "Panels and Cohorts: Design and Model in the Study of Voting Turnout." *Sociological Methodology* 18:1–35.

- Easterlin, Richard. 1978. "What Will 1984 Be Like? Socioeconomic Implications of Recent Twists in Age Structure." *Demography* 15:397-432.
- . 1980. *Birth and Fortune: The Impact of Numbers on Personal Welfare*. New York: Basic.
- England, Paula, George Farkas, Barbara S. Kilbourne, and Thomas Dou. 1988. "Explaining Occupational Sex Segregation and Wages. Findings from a Model with Fixed Effects." *American Sociological Review* 53:544-58.
- Firebaugh, Glenn. 1989. "Methods for Estimating Cohort Replacement Effects." *Sociological Methodology* 19:243-62.
- Firebaugh, Glenn, and Kenneth E. Davis. 1988. "Trends in Antiblack Prejudice, 1972-1984: Region and Cohort Effects." *American Journal of Sociology* 94:251-72.
- Firebaugh, Glenn, and Brian Harley. 1991. "Trends in U.S. Church Attendance: Secularization and Revival, or Merely Life Cycle Effects?" *Journal for the Scientific Study of Religion* 30:487-500.
- Glenn, Norval D. 1976. "Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period, and Cohort Effects." *American Sociological Review* 41:900-904.
- . 1980. "Values, Attitudes, and Beliefs." Pp. 596-640 in *Constancy and Change in Human Development*, edited by Orville Brim and Jerome Kagan. Cambridge, Mass.: Harvard University Press.
- . 1989. "A Caution about Mechanical Solutions to the Identification Problem in Cohort Analysis: Comment on Sasaki and Suzuki." *American Journal of Sociology* 95:754-61.
- Hobcraft, John, Jane Menken, and Samuel Preston. 1982. "Age, Period, and Cohort Effects in Demography: A Review." *Population Index* 48:4-43.
- Hout, Michael, and David Knoke. 1975. "Change in Voting Turnout, 1952-1972." *Public Opinion Quarterly* 39:52-68.
- Inglehart, Ronald. 1971. "The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies." *American Political Science Review* 65:991-1017.
- . 1985. "Aggregate Stability and Individual-Level Flux in Mass Belief Systems: The Level of Analysis Paradox." *American Political Science Review* 79:97-116.
- Jennings, M. Kent. 1987. "Residues of a Movement: The Aging of the American Protest Generation." *American Political Science Review* 81:367-82.
- Katosh, J. P., and M. W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures." *Public Opinion Quarterly* 45:519-35.
- Key, V. O., Jr. 1949. *Southern Politics in State and Nation*. New York: Vintage.
- Kleppner, Paul. 1982. "Were Women to Blame? Female Suffrage and Voter Turnout." *Journal of Interdisciplinary History* 12:621-43.
- Koslowski, Karl. 1986. "Isolating Age, Period, and Cohort Effects in Developmental Research." *Research on Aging* 8:46-79.
- Lerner, Richard L. 1984. *On the Nature of Human Plasticity*. New York: Cambridge University Press.
- Lesthaeghe, Ron, and Johan Surkyn. 1988. "Cultural Dynamics and Economic Theories of Fertility Change." *Population and Development Review* 14:1-45.
- Liker, Jeffrey K., Sue Augustyniak, and Greg J. Duncan. 1985. "Panel Data and Models of Change. A Comparison of First Difference and Conventional Two-Wave Panels." *Social Science Research* 14:80-101.
- Mannheim, Karl. (1927) 1952. "The Problem of Generations." Pp. 276-322 in *Essays on the Sociology of Knowledge*. London: Routledge & Kegan Paul.
- Merriam, C. E., and H. F. Gosnell. 1924. *Non-voting: Causes and Methods of Control*. Chicago: University of Chicago.
- NES (National Election Studies). 1990. *American National Election Studies: Cumulative Data File, 1952-1988*. Inter-university Consortium for Political and Social Research, Ann Arbor, Mich.



- Norpoth, Helmut. 1987. "Under Way and Here to Stay: Party Realignment in the 1980s?" *Public Opinion Quarterly* 51:376-91.
- Ortega y Gasset, Jose. 1933. *The Modern Theme*. New York: Norton.
- Petrocik, John R. 1991. "An Algorithm for Estimating Turnout as a Guide to Predicting Elections." *Public Opinion Quarterly* 55:643-47.
- Presser, Stanley, and Michael Traugott. 1992. "Little White Lies and Social Science Models: Correlated Response Errors in a Panel Study of Voting." *Public Opinion Quarterly* 56:77-86.
- Roberts, Carl W. 1986. "Tracing Formative Influences on Event Recall: A Test of Mannheim's Sensitivity Hypothesis." *Social Forces* 65:74-86.
- Rosow, Irving. 1978. "What Is a Cohort and Why?" *Human Development* 21:65-75.
- Rusk, Jerrold G., and John Stucker. 1978. "The Effect of the Southern System of Election Laws on Voting Participation: A Reply to V. O. Key, Jr." Pp. 198-250 in *The History of American Electoral Behavior*, edited by Joel Silbey, Allan Bogue, and William Flanigan. Princeton, N.J.: Princeton University Press.
- Ryder, Norman B. 1965. "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30:843-61.
- Sasaki, Masamichi, and Tatsuzo Suzuki. 1987. "Changes in Religious Commitment in the United States, Holland, and Japan." *American Journal of Sociology* 92: 1055-76.
- Schuman, Howard, and Jacqueline Scott. 1989. "Generations and Collective Memories." *American Sociological Review* 54:359-81.
- Schuman, Howard, and Cheryl Rieger. 1992. "Historical Analogy, Generational Effects, and Attitudes toward War." *American Sociological Review* 57:315-26.
- Sears, David O. 1990. "Whither Political Socialization Research? The Question of Persistence." Pp. 69-97 in *Political Socialization, Citizenship Education, and Democracy*, edited by O. Ichilov. New York: Teachers College Press.
- Sigelman, L. 1982. "The Nonvoting Voter in Voter Research." *American Journal of Political Science* 26:47-56.
- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80:613-24.
- Stinchcombe, Arthur. 1984. "Why We Have No Theory of Cultural Change." Paper presented at the annual meeting of the American Sociological Association, San Antonio.
- Weil, Frederick D. 1987. "Cohorts, Regimes, and the Legitimation of Democracy: West Germany since 1945." *American Sociological Review* 52:308-24.
- Welch, Finis. 1979. "Effects of Cohort Size on Earnings. The Baby Boom Babies' Financial Bust." *Journal of Political Economy* 87:S65-S97.
- Welch, Susan. 1977. "Women as Political Animals? A Test of Some Explanations of Male-Female Political Participation Differences." *American Journal of Political Science* 21:711-30.
- Wilmoth, John R. 1990. "Variation in Vital Rates by Age, Period, and Cohort." *Sociological Methodology* 20:295-336.
- Wolfinger, R. E., and S. J. Rosenstone. 1980. *Who Votes?* New Haven, Conn.: Yale University Press.

# Class Mobility and Political Preferences: Individual and Contextual Effects<sup>1</sup>

Nan Dirk De Graaf and Paul Nieuwbeerta  
*University of Nijmegen*

Anthony Heath  
*Nuffield College*

The authors test several hypotheses about the impact of intergenerational class mobility on political party preferences. Tests using cross-national data sets representing Britain, the Netherlands, Germany, and the United States over the period 1964–90 suggest a process of acculturation to the class of destination. The authors hypothesized that a class with a high degree of demographic identity influences newcomers more than a class with low demographic identity does and that, the more left-wing inflow there is into a class, the more likely the immobile members are to have left-wing political preferences. The data did not confirm these hypotheses. A macro analysis does, however, show that the level of class voting is weakened by a compositional mobility effect.

## INTRODUCTION

It is generally true that a member of the working class is more likely to vote for a left-wing political party than is a member of the middle class, although the extent of such class voting varies greatly between countries. In Britain, for example, it is relatively strong, whereas in America, particularly in recent years, it has been rather weak. Moreover, even in countries where class voting is relatively strong, the proportion voting for the “natural” party of their class rarely exceeds 60% (Heath et al. 1991, pp. 68–69).

<sup>1</sup> This research was supported by the Royal Netherlands' Academy of Arts and Sciences, the Netherlands Organization for Scientific Research (grant no. 500-279-001), and the Economic and Social Research Council (award Y303253001). We especially wish to thank Wout Ultee and three *AJS* reviewers for their detailed comments, Harry Ganzeboom, Herbert Kritzer, Mildred Schwartz, and Marco Steenbergen for their suggestions, and Jeroen Weesie and David Firth for their statistical advice. Direct correspondence to Nan Dirk De Graaf, University of Nijmegen, Department of Sociology, P.O. Box 9104, 6500 HE Nijmegen, Netherlands E-mail: U211313@vm.uci.kun.nl

To understand fully the relationship between class position and political preference one has to remember that socioeconomic classes in the advanced industrial societies are far from homogeneous and static. In all Western countries, although not always to the same extent, people move up and down the social ladder with respect to their parents. Processes of social mobility, therefore, may help to explain "class-deviant" behavior within a country (Abramson 1972).

A number of empirical studies have investigated the impact of social mobility on voting behavior (Lipset and Zetterberg 1956; Lipset and Bendix 1959; Barber 1970; Abramson 1972; Kelley 1992; Turner 1992; Weakliem 1992). Most of these adopt a microsociological perspective and focus on the impact of an individual's mobility experience on his or her voting behavior; almost invariably they have shown that the political behavior of the mobile is intermediate between that of the stable members of their origin and destination classes. However, there is also an older tradition of macrosociology that looks at the emergent properties of social formations (Durkheim 1897; Blau 1964; Blau and Schwartz 1984). That is to say, rates of social mobility may affect the degree of cohesion and solidarity within a particular class, and this in turn may influence the overall level of class voting.

Mobility processes may therefore have important consequences for levels of class polarization, and they may help to explain both over-time and cross-national variations in class voting (De Graaf and Ultee 1990; Heath et al. 1991; Nieuwbeerta 1995). In particular higher absolute rates of mobility may be expected to reduce class polarization by increasing the proportion of mobile voters with intermediate propensities to support the left (which we term compositional effects) and also by weakening class solidarity and pulling even the nonmobile voters toward the center (which we term the contextual effects on the immobile).

In this article we first discuss two main theories of the relation between class and political preference. On the basis of these theories we suggest some micro hypotheses about the effects of an individual's mobility experiences on political preferences. Next, we consider macro hypotheses about the compositional and contextual effects of mobility rates. In the next four sections, we describe our data, discuss our models, and present the results of testing our micro and macro hypotheses. In these tests we restrict ourselves to the analysis of men. De Graaf and Heath (1992) have shown that working women should not be assigned to social classes solely on the basis of their husband's occupation or solely on the basis of their own occupation. For women, more complex models are required that take account of the interactions between husbands' and wives' occupations. We therefore believe that women's mobility deserves a separate

treatment in its own right, and we hope to present this in a separate publication.

#### MICRO EFFECTS OF INTERGENERATIONAL MOBILITY ON POLITICAL PREFERENCE

In order to obtain a better insight into the consequences of intergenerational class mobility for political preference, we begin with two main theories of individual voting behavior. The first is the economic theory of political behavior (Downs 1957), sometimes also known as the "instrumental" theory. The basic idea here is that voting behavior is rational and self-interested: People vote for the party whose policies will bring them the greatest utility in the future. Class voting can then be explained on the grounds that people in lower social classes have an interest in redistributive policies, which are typically espoused by left-wing parties, while the members of higher social classes have an interest in opposing such policies (Lipset et al. 1954; Converse 1958).

The second theory, labeled by Heath, Jowell, and Curtice (1985, p. 9) as the "expressive" theory, perceives voting as a social act rather than an instrumental act. The assumption of this theory is that political identities are developed through interaction with others. Voting behavior is thus an expression of a political identity and will in turn reflect the norms and values of one's normative reference group. With this theory, too, it is rather straightforward to explain the relationship between class position and political preference. In many cases a person associates with other people occupying the same class position. They are raised by them, live in the same area, attend the same school, work together as colleagues, and thus learn the traditional culture of their shared class. Consequently, they will vote as their class members traditionally vote.

Although these two theories are quite different with respect to their initial assumptions, in predicting the relation between class and political preference, they are not contradictory but complementary (Heath et al. 1985, p. 9). To phrase the complementary character of both theories: People can vote for the same party both because of their mutual interest and because they are influenced by each other. We might add that associating with others from the same class may make people more conscious of their common interests and of the party that serves their interests best.

However, the two theories do have rather different implications for the relationship between social mobility and political preference. In its simplest version, the instrumental theory predicts that the political preferences of the mobile will be identical to those of their class of destination, although, given the evidence on counter-mobility (Girod 1971; Goldthorpe

1980), one might expect there to be some modest influence from social origins too. That is, some people might expect their current class positions to be temporary and might anticipate returning to their class of origin; hence they might define their long-run interests as those of their class of origin. On the basis of the instrumental theory, therefore, one would predict that *the political preferences of the mobile will be closer to the typical political preferences of their class of destination than to that of their class of origin*. This is what De Graaf and Ultee (1990) called the "economic hypothesis," and it is our first hypothesis at the micro level.

The expressive theory, on the other hand, predicts a larger role for social origins, since the culture of one's origin class is likely to be particularly important in early political socialization. However, the older someone is, the more distant is their primary socialization in their family and class of origin and, conversely, older respondents will in general have had a longer period in their class of destination (most intergenerational class mobility between the broad classes that we have identified takes place at a relatively early stage of the career and little occurs after the age of 35; Goldthorpe 1980, pp. 69–71). Our "acculturation hypothesis" therefore states that *the older one is the more the impact of the class of origin diminishes relative to that of the class of destination*. This hypothesis is in line with Blau's "pattern of acculturation" (1956) and follows directly from the expressive theory. It is unlikely that the mobile will have no social contacts with their class of origin. Social networks change slowly, and it may take time for the mobile to integrate socially into their class of destination and to lose their old social identity.

A further issue that has often been raised is whether there is an asymmetry in the patterns of adaptation of upwardly and downwardly mobile people. Lipset (1960), for example, assumed that upwardly mobile people adapt more quickly to their destination class than do those who are downwardly mobile. His underlying idea was that people in general prefer to adopt the more prestigious identity and, thereby, to maximize their status. We term this the status maximization hypothesis. That is, people may prefer to take as their normative reference group whichever is the higher of their classes of origin and destination. The "status maximization hypothesis," then, is that *downwardly mobile persons orient themselves more to their origin class, while upwardly mobile persons will orient themselves more to their destination class*.<sup>2</sup> Similar statements

<sup>2</sup> We would like to emphasize that this hypothesis is not just the prediction that political preference will be a weighted average between origin and destination. The literature on social mobility and political preference often suggests that, due to the "shock of mobility" (comparable to the often assumed status inconsistency effect), mobility has an extra independent effect. Turner (1992) labels this the "effect of

have been made by Parkin (1971, pp. 51, 54), Knoke (1973), and Thorburn (1979), although subsequent empirical research has generally failed to confirm this hypothesis (Abramson 1972; De Graaf and Ultee 1990; Nieuwebeerta and De Graaf 1993; Clifford and Heath 1993).

#### MACRO EFFECTS OF INTERGENERATIONAL MOBILITY ON POLITICAL PREFERENCE

The theoretical questions raised at the macrolevel include some of the central ones of stratification research, for example, questions about the potential for class-based collective action. The classic statement of this was Marx's explanation for the absence of social classes in America. Thus in a well-known passage, Marx (1926, p. 33) argued that in America "classes are not yet fixed, but in continual flux, with a persistent interchange of their elements." In other words, high rates of social mobility undermine class formation. Sorokin (1957, p. 538) argued that social mobility facilitates atomization and diffusion of solidarity and antagonisms. Similar ideas have been widespread in the many accounts of the decline of class in contemporary societies: Increasing rates of social mobility have, it is claimed, tended to weaken the cohesion of the classes and have been one of the factors in class dealignment (Clark and Lipset 1991).

Goldthorpe (1980) has followed up these broad claims with more specific hypotheses in his work on class formation. He distinguishes between the demographic identity of a class (defined by its patterns of inter- and intragenerational mobility) and the sociopolitical orientations that demography may promote. The implications of these demographic considerations for normative patterns seem clear enough. As Goldthorpe suggests (1980, p. 268), "the fact that over recent decades [the service class] has recruited from very diverse sources must have seriously reduced its sociocultural distinctiveness; to some extent mobile men carry with them the normative and relational patterns of their class origins into their class of destination, and thereby increase its internal heterogeneity. However, the case of the working class is clearly a different one. It is from its recruitment patterns—that is from the homogeneity of the social origins of its members—that its demographic identity primarily derives; and in turn, one may then suppose, its sociocultural homogeneity is in this way also favoured."

There are two rather different ways in which one could interpret these ideas and then put them to the test—the compositional and the context-

---

mobility *per se*." We feel that our status maximization hypothesis is a more specific hypothesis on mobility effects.

tual. First, within an individualistic perspective, one might claim that the sociocultural orientations of a class are simply the summation of those of the individuals that make up the class. We term this the "compositional interpretation." As the service class expands and includes more people from non-service-class origins, so its demographic homogeneity will decline. If these newcomers retain to some extent the political orientations of their class origins, and if these orientations are different from those of the stable members of the class of destination, it follows arithmetically that the political complexion of the class as a whole must change. However, it is important to acknowledge that it is not just the amount of inflow mobility that may be important. The impact of inflow mobility may also depend on the political preferences of the newcomers. For example, we know that the political preferences of the service class are not very different from the preferences of the routine nonmanual class whereas the manual class differs substantially in this respect. This suggests that inflow into the service class from the routine nonmanual class may have less impact politically than the same amount of inflow from the manual class. The compositional effects of mobility depend, therefore, both on the amount and on the political character of that mobility.

However, we might also wish to argue, on Durkheimian lines, that the culture of a class is more than the sum of individual orientations. In other words, class culture may be an emergent property of the class, deriving from the interaction between the individuals who currently make up the class. Thus a class with a high degree of demographic identity may have a stronger normative subculture than one with a lower degree of demographic identity, and this in turn may exert a stronger pressure on newcomers to the class. The process of acculturation to the norms and values of the class of destination may be faster and more complete if the class has a high degree of demographic identity. We term this the "contextual interpretation." Weakliem (1992, p. 155) neatly summarized this argument as follows: "Effects of origin and destination class may vary according to the nature of socialization of different classes. Some social groups (classes) may have a strong culture which places a definite stamp on newcomers. . . . Other groups (classes) may have little sense of common identity and consequently little impact on the views of new arrivals." Our "contextual hypothesis for the mobile," therefore, states that *a class with a low level of inflow mobility (and therefore a high level of demographic identity) will have a greater impact on newcomers than will a class with a higher rate of inflow mobility.*<sup>3</sup>

<sup>3</sup> A similar hypothesis was postulated by Ultee and De Graaf (1991) to explain culture consumption.

So far we have considered only hypotheses about the behavior of inter-generationally mobile individuals. Although most of the literature concentrates on the mobile, this is only half the story. As Blau and Schwartz (1984, p. 55) have pointed out, immobile people too can be affected by mobility processes: "Mobility has cumulative effects on intergroup processes. First, since mobile persons have more intergroup associates than nonmobile persons, many mobiles raise the group's average [our compositional effect]. Second, they sometimes introduce their friends from the two groups, initiating intergroup relations between these. Third, the consequent high rates of intergroup relations in communities with many mobiles indicate weaker ingroup pressures [our contextual hypothesis for the mobile]. Finally, the relatively frequent intergroup relations lessen ingroup salience, even among those who themselves are not involved in intergroup associations, further facilitating contacts in the future. The inference is that extensive mobility promotes intergroup relations of a population's nonmobile as well as its mobile members."

Thus if there is a large inflow into a class, the chances that an immobile class member will be influenced by the newcomers may be relatively high (see also Lipset 1960). The impact of the newcomers may be expected to depend on the political orientations that they bring with them, and so our "contextual hypothesis for the immobile" is that *the more left-wing mobility into a class there is, the more likely are the immobile members of that class to have a left-wing political preference.*

#### DATA

While the micro hypotheses can be tested on a single data set from one country, it is of some interest to test whether they hold true for other countries or at other times. In the case of the macro hypotheses, moreover, it is essential to increase the number of data sets. In our contextual hypotheses the unit of analysis becomes the class and, with the six-class schema that we use in this article, each data set thus yields six observations. In order to obtain an adequate number of observations we therefore employ a total of 55 data sets from four countries: Britain, Germany, the Netherlands, and the United States. The data file for the Netherlands contains 13 Dutch representative surveys over the period 1970–90. The U.S. file consists of 17 GSS data sets over the period 1972–90. The British file consists of seven National Election Studies covering the 1964–87 period. For Germany the file is built upon 19 representative surveys covering the period 1969–90. The surveys used are listed in the data references list. The merged data files contain comparable information on respondent's class, father's class, and respondent's poli-



tical preference. Additional comparable relevant variables that are included in our data set are age, religion, ethnicity, and year of interview.

From the data, we have selected male respondents: 11,949 for the United States, 8,349 for Britain, 10,489 for the Netherlands, and 19,216 for Germany. From these respondents, we selected those 18 years old or older who had a valid score on all relevant variables. These restrictions reduce the number of cases to 9,372 for the United States, 6,168 for Britain, 7,086 for the Netherlands, and 10,596 for Germany.

### Political Preference

Political preference, our dependent variable, is measured in somewhat different ways in the different surveys. In all cases in the British and U.S. surveys (and in many of the German and Dutch surveys) respondents were asked to name the political party they voted for in the most recent national election. In the other German and Dutch surveys the questions tapped vote intention.<sup>4</sup> In this article we use the phrase "political preference" for all these measures. The difference between these measures of political preference is a limitation that must be borne in mind. However, experiments with the different treatments of the dependent variable showed that our conclusions are not sensitive to these differences. In order to achieve a measure of international comparability we have followed Bartolini and Mair's (1990) and Franklin, Mackie, and Valen's (1992) procedures and dichotomized political preference into a preference for left-wing parties (coded "1") versus other parties (coded "0").<sup>5</sup>

<sup>4</sup> Surveys (see data references list) asked about voting preferences in the following ways: which party the respondent voted for in the last national elections (Barnes and Kaase 1976; Werkgroep Nationaal Kiezersonderzoek 1977; Van der Eijk, Niemoeller, and Eggen 1981; Van der Eijk et al. 1982; Van der Eijk, Irwin, and Niemoeller 1986; ZUMA 1980; International Social Science Program 1987; Butler and Stokes 1963-66, 1970; Crewe, Saerlvik, and Alt 1974; Crewe, Saerlvik, and Robertson 1979; Heath et al. 1983; Heath et al. 1989; Davis and Smith 1991); which party the respondent would vote for if national elections were held today/tomorrow (Heunks et al. 1973; Verba, Nie, and Kim 1971; Allerbeck et al. 1980; Felling, Peters, and Schreuder 1985; Eisinga et al. 1990; Barnes and Kaase 1976; ZUMA 1988, 1990); which national political party the respondent prefers (Hermkens and Van Wijngaarden 1976; CBS 1977; Arts et al. 1989); and which party the respondent expected to vote for in the upcoming national elections (Kaase and Schleth 1969; Klingemann and Pappi 1969; Allerbeck et al. 1980). Exact wording for the questions used in the surveys can be found in the original codebooks.

<sup>5</sup> The following parties are coded as left wing: Germany—Social Democrats, Communist Party, Independent Social Democrats, and Action for Democratic Progress; Netherlands—Labour Party, Social Democratic League, Socialist Party, Revolutionary

TABLE 1  
SOCIAL CLASS SCHEMA

Number	Title	Description	EGP Categories*
1 .....	Service class	Professionals, administrators, and managers; higher-grade technicians; supervisors of nonmanual workers	1, 2
2 . . . . .	Routine nonmanual workers	Routine nonmanual employees in administration and commerce; sales personnel; other rank-and-file service workers	3
3 .....	Petty bourgeois	Small proprietors and artisans, with and without employees	4a, 4b
4 .....	Farmers	Farmers, smallholders, and other self-employed workers in primary production	4c
5 .....	Skilled and unskilled manual workers	Lower-grade technicians; supervisors of manual workers; skilled manual workers; semi- and unskilled, nonagricultural manual workers	5, 6, 7a
6 .....	Agricultural laborers	Agricultural and other workers in primary production	7b

\*Numbers indicate the categories of the EGP schema (Erikson et al 1979; Erikson and Goldthorpe 1992) that we have combined to define our categories.

### Social Class

The social class variable is coded according to a collapsed version of Erikson, Goldthorpe, and Portocarero's (EGP schema; see Erikson, Goldthorpe, and Portocarero 1979; Erikson and Goldthorpe 1992). We use this schema because it has proved to be useful in comparative intergenerational mobility research (Ganzeboom, Luijkx, and Treiman 1989) and in analyses of the consequences of social mobility for political preference (Nieuwbeerta and De Graaf 1993; Clifford and Heath 1993). Because of limitations in the data (e.g., because in some surveys detailed information on self-employment and supervisory status is lacking), we are not able to use more extended versions of the EGP schema.

The class schema that we shall use is shown in table 1. Respondents

Socialist Party, Communist Party, Pacifist Socialist Party, and Democratic Socialist '70; Great Britain—Labour Party; United States—Democratic Party.

were coded into classes based on their occupation, self-employment, and supervisory status.<sup>6</sup> We should note that the schema is not intended to be strictly hierarchical and that it has a nonlinear relationship with vote.

### Age

We use AGE as our explanatory variable for testing the acculturation hypothesis. This continuous variable is coded from -22 (18 years old, our youngest respondents) to 54 (94 years old, our oldest respondents). The advantage of this linear transformation is that the score "0" is about the average age in each country. We also need to take account of the main effects of age on party preference. For this we include three age groups (18-30, 31-50, and 51 and older) as covariates, since there appeared to be typically nonlinear relationships between age and party preference.

### Other Variables

Although the impact of class mobility is our main theoretical concern, we must also take into account the impact of other relevant variables. Religion is particularly relevant, since we know, for example, that in Germany and the Netherlands Catholics tend to vote for right-wing parties, while there are religious differences in political behavior in Britain and the United States (Heath, Taylor, and Toka 1993). We divided church membership into five groups: Catholic, Protestant (Reformed), Orthodox Reformed (for the Netherlands only), other denomination, and no religion.

Given the importance of ethnicity for predicting political preference in the United States, we include ethnicity (0 = nonblack; 1 = black) for the United States only. For the other three countries there were no minority ethnic groups large enough that ethnicity could sensibly be included as an independent variable.

A final important control variable is the year of interview, which we recoded into a set of dummies representing several periods: 1961-65, 1966-70, 1971-75, 1976-80, 1981-85, and 1986-90. We clearly need to control for the varying popularity of the political parties over the years.

<sup>6</sup> Coding the respondents into the EGP scheme takes two steps. First, we recode the original occupation codes into the International Standard Classification of Occupation (ISCO) codes (ILO 1969). Second, we translate these ISCO codes into EGP scores through the Ganzeboom, Luijkx, and Treiman (1989) recoding schema.

## HOW TO MODEL MICRO EFFECTS OF INTERGENERATIONAL MOBILITY

Our hypotheses distinguish between the political preferences of mobile and immobile individuals, and we therefore need a statistical model that corresponds to these substantive concerns. Conventional approaches that model political behavior as a function of origin and destination class will not be appropriate since the explanatory categories will include mixtures of mobile and immobile respondents. A substantively more appropriate model, in our view, is Sobel's (1981, 1985) diagonal reference model,<sup>7</sup> in which the behavior of mobile respondents is modeled as a function of the behavior of the immobile respondents in the classes of origin and destination. The model therefore allows us to test specifically whether the mobile are closer in their political behavior to the immobile members of their class of destination or to the immobile members of their class of origin.<sup>8</sup>

The theoretical importance of taking the immobile as the reference has been suggested by Sorokin (1957, pp. 509–10), who argued that "if we want to know the characteristic attitudes of a farmer, we do not go to a man who has been a farmer for a few months, but go to a farmer who is a farmer for life." Even better, we would argue, go to a farmer who was born and bred a farmer. We can think of the immobile respondents as representing the core of the class and defining its norms and values to which newcomers may or may not acculturate.

Sobel's original model was designed for predicting a metric dependent variable. Given our binomial dependent variable, we fit instead of Sobel's model a logistic version. The logistic version can be expressed as follows:

$$\text{prob}(Y_{ijk} = 1) = 1/(1 + e^{-\ln}); \quad (1)$$

$$\ln = p\alpha_i + (1 - p)\alpha_j + \beta_L \text{cov}_L, \quad (2)$$

where  $Y_{ijk} = 1$  if respondent  $k$  in the  $ij$ th cell of the mobility table has a left-wing preference and 0 if not. Subscript  $j$  stands for father's class and  $i$  for respondent's class. There is one parameter,  $\alpha$ , for each diagonal cell, representing the expected mean behavior of the core (the stable members) of each class. Parameters  $p$  and  $(1 - p)$  are destination and origin weights that indicate the relative importance of the core destination

<sup>7</sup> Models of this type were originally known as diagonal mobility models. De Graaf and Ganzeboom (1990) argued that the label "diagonal reference model" is more appropriate.

<sup>8</sup> For a comparison of diagonal models with the conventional ones see Hendrickx et al. (1993).

and origin classes.<sup>9</sup> In equation (2) we estimate one  $\beta$  coefficient for each covariate (cov). Consequently, if we have  $L$  covariates, we have to estimate  $L$   $\beta$  parameters.<sup>10</sup>

# RESULTS: TESTS OF THE MICRO HYPOTHESES

Table 2 presents the proportion of respondents with a left-wing political preference in each category of the intergenerational mobility table in each country (the different survey years being combined). We see that class has a clear nonlinear relationship with political preference. The petty bourgeois and farmers are by far the least likely to have a left-wing party preference in the Netherlands, Germany, and Britain, while in all cases the manual workers have the highest probability of a left-wing preference. Farmers in America are the most notable exception to these general propositions, and while there are similarities in the relationship between class and political preferences in all four countries, it has to be said that the similarity between the three European countries is particularly striking. The anomalous voting behavior of farmers in the United States is probably to be explained by the fact that the Democratic Party in the United States, unlike the left-wing parties in our three European countries, has traditionally endorsed farm-support measures.

The size of the class differences in left-wing politics, do, however, differ considerably. We can use the index of dissimilarity as an overall measure of class differences in voting behavior.<sup>11</sup> The index proves to be

<sup>9</sup> We assume the restriction that  $0 \leq p \leq 1$ . The parameter  $p$  does not have this restriction in the estimating procedure, however, because, in order to get the best parameter estimate in the iterative procedure,  $p$  might get higher than 1 or lower than 0. The diagonal reference models do not offer the appropriate design when  $p$  does not fit in the 0-1 interval, but this situation did not arise in any of the models used.

<sup>10</sup> Sobel's (1985) more general model allows a separate  $\beta$  coefficient for each covariate for each diagonal reference cell  $i$ . The logistic version is

$$\text{prob}(Y_{ijt} = 1) = 1/(1 + e^{-\lambda_{ijt}}); \quad (1)$$

$$\lambda_{ijt} = p(\alpha_i + \beta_L \text{cov}_L) + (1 - p)(\alpha_j + \beta_L \text{cov}_L). \quad (3)$$

We have, however, no reason to assume a different relationship between our covariates and political preference for each diagonal cell. The advantage is that we have a more parsimonious model and can therefore rewrite eq. (3) into the simpler eq. (2).

<sup>11</sup> The usual method of measuring the level of class voting is the Alford index, but this index can be used only when the classes have been dichotomized. The index of dissimilarity does not require any restriction on the number of classes. The index is obtained by summing the differences between the two distributions (i.e., between the class distribution of left-wing voters and the class distribution of right-wing voters) and dividing by 2. It thus measures the proportion who would have to change their political preferences to make the two distributions identical. The calculations in the text refer to the overall class distributions, including both mobile and immobile respondents, and have been calculated directly from table 2.

rather low in the United States at 15 and highest in Britain at 36, with the Netherlands (23) and Germany (22) in between.

Table 2 also shows, as previous research has done, that the political preferences of the mobile typically differ from those of the stable members of their class of destination. For example, in almost all cases individuals who are downwardly mobile into the manual working class display a lower level of left-wing political preference than do the stable members of the manual working class. Conversely, individuals who are mobile into the petty bourgeois tend to display a higher level of left-wing political preference than do the intergenerationally stable petty bourgeois.

We now move on to formal modeling of these data and tests of our hypotheses. To do this we combine the four national data files. We first fit baseline model A, including one parameter for the destination and origin weights ( $p$  and  $[1 - p]$ ) and separate parameters for the religion, age, period,<sup>12</sup> and ethnicity covariates. In this model we thus fit a common destination/origin weight but allow the covariates and the class parameters to differ in the four countries. The model therefore uses 63 parameters.<sup>13</sup> A more parsimonious model that fitted common parameters for the classes and covariates could be explored, but since our primary interest is in the effects of mobility, a conservative strategy toward these covariates is appropriate.

The reason for the country-specific modeling of the class parameters is that we cannot assume the socioeconomic situation of a class to be identical in the four countries. Furthermore, the left-wing and right-wing political parties in our four countries may not be equally left- or right-wing in their political stances and may not equally espouse perceived class interests. For example, the Labour Party in Britain is substantially more left-wing than the Democratic Party in the United States.<sup>14</sup> Since

<sup>12</sup> We also fitted a dummy variable for each year instead of for each period. Using a dummy for each year, however, results in a large number of parameters. This is especially troublesome when we want to test whether the differences between the countries are significant. Since reducing these parameters by applying five-year periods did not change our results significantly, we preferred to apply these periods.

<sup>13</sup> We use one parameter for the weight, six parameters in each of the four countries for the diagonal reference categories, three parameters for religion in Britain, Germany, and the United States, and four for the Netherlands, two parameters for the age categories in each country, one for ethnicity in the United States, five for period in Britain, four for Germany and the Netherlands (since the 1961–65 period is not available in these countries), and three for the United States (where neither the 1961–65 nor 1966–70 periods are available).

<sup>14</sup> In both the U.S. and British National Election Studies of 1992, identical questions were asked about whether “the government should see to it that every person has a job and a good standard of living . . . [or] . . . should just let each person get ahead on their own.” Respondents were asked for their own positions (on a seven-point scale) and for their perceptions of where the parties stood. In Britain 63% of respondents placed Labour on points 1 or 2 of the scale (the most left-wing positions) whereas in America only 30% placed the Democrats on points 1 or 2.

TABLE 2  
PERCENTAGE OF LEFT-WING VOTERS BY RESPONDENT'S AND FATHER'S CLASS

FATHER'S CLASS	RESPONDENT'S CLASS						All Workers
	Service Class	Routine Nonmanual	Petty Bourgeois	Farmers	Manual Workers	Agricultural Workers	
Germany:							
Service class .....	32 (33)	34 (17)	22 (12)	33 (9)	46 (8)	11 (7)	35
Routine nonmanual ....	40 (10)	45 (11)	17 (6)	0 (1)	57 (5)	67 (5)	46
Petty bourgeois .. .....	28 (10)	28 (11)	12 (33)	0 (2)	37 (6)	52 (2)	28
Farmers.....	21 (8)	19 (8)	15 (11)	7 (89)	37 (12)	18 (26)	25
Manual workers ....	46 (38)	46 (50)	16 (38)	22 (5)	61 (65)	52 (35)	54
Agricultural workers ..	36 (1)	52 (3)	21 (1)	0 (2)	60 (4)	52 (25)	55
All.....	37 (100)	40 (100)	16 (100)	8 (100)	55 (100)	41 (100)	44
N .....	3,146	1,250	484	345	5,245	126	10,596
Great Britain:							
Service class .....	13 (32)	22 (21)	14 (13)	9 (10)	35 (6)	10 (10)	19
Routine nonmanual ...	15 (8)	33 (11)	25 (5)	0 (0)	42 (3)	0 (0)	26
Petty bourgeois .....	14 (10)	20 (9)	11 (24)	0 (9)	31 (6)	0 (4)	20
Farmers....	16 (3)	7 (3)	25 (6)	8 (67)	50 (3)	24 (13)	27
Manual workers .....	27 (46)	39 (54)	21 (49)	0 (10)	61 (76)	40 (26)	50
Agricultural workers ..	15 (2)	34 (3)	34 (3)	58 (4)	55 (5)	41 (47)	47
All.....	20 (100)	32 (100)	19 (100)	9 (100)	56 (100)	34 (100)	40
N .....	1,725	507	483	113	3,244	96	6,168

Netherlands.									
Service class	...	21 (36)	26 (23)	24 (11)	26 (3)	38 (10)	22 (6)	25	
Routine nonmanual	.	28 (13)	34 (15)	13 (4)	50 (1)	46 (6)	33 (2)	33	
Petty bourgeois	.....	19 (13)	28 (14)	7 (45)	6 (4)	37 (11)	25 (5)	23	
Farmers.....		18 (10)	16 (8)	11 (14)	5 (84)	27 (12)	12 (37)	16	
Manual workers.....		37 (26)	39 (37)	30 (21)	30 (4)	53 (55)	45 (20)	46	
Agricultural workers..		31 (3)	43 (4)	23 (5)	32 (5)	51 (7)	37 (30)	42	
All.....		26 (100)	32 (100)	15 (100)	8 (100)	46 (100)	28 (100)	33	
N .....		2,443	1,183	390	404	2,511	155	7,086	
United States:									
Service class	.....	31 (38)	31 (32)	23 (21)	9 (3)	40 (15)	14 (12)	33	
Routine nonmanual...		35 (7)	21 (10)	49 (3)	0 (1)	34 (4)	0 (0)	33	
Petty bourgeois...		33 (7)	54 (7)	39 (15)	54 (2)	43 (7)	43 (4)	41	
Farmers .....		33 (12)	39 (10)	39 (18)	34 (84)	48 (19)	43 (46)	41	
Manual workers.	.	35 (36)	42 (41)	36 (42)	38 (8)	49 (53)	29 (24)	43	
Agricultural workers..		40 (1)	95 (1)	73 (2)	72 (2)	60 (3)	63 (14)	62	
All ..		33 (100)	38 (100)	35 (100)	35 (100)	46 (100)	39 (100)	40	
N ....		3,148	930	447	367	4,393	87	9,372	

Note: Percentages of inflow mobility are given in parentheses



the level of class voting is likely to depend on the congruence between perceived class interests and perceived party policies, we must clearly take into account relevant national differences. By allowing the diagonal parameters to differ between countries, we control for these differences. Similar considerations apply to the covariates and, for example, to the well-known national differences in the role of religion in voting behavior.

In this baseline model, the comparison of the origin and destination weights allows us to assess our first hypothesis at the microlevel, namely the economic hypothesis. In table 3 we report the fit ( $G^2 = -2 \log$  likelihood) of this model. The quantity  $-2 \log$  likelihood does not have a straightforward interpretation, but the goodness of fit of one model relative to another can be assessed using the fact that, provided the two models in question are nested, the difference has a chi-square distribution with degrees of freedom equal to the difference between the two models in the number of their parameters.

TABLE 3  
NESTED LOGISTIC DIAGONAL REFERENCE MODELS FOR THE RELATIVE  
INFLUENCE OF RESPONDENT'S CLASS AND FATHER'S CLASS  
ON RESPONDENT'S POLITICAL PREFERENCE

Model and Description	$G^2$	BIC	$df$ Used
A—Baseline .....	39,194.3	-306,023.0	63
	$-\Delta G^2$	$\Delta BIC$	$\Delta df$ Used
B—Acculturation . . . . .	39.8	-29.4	1
C—Status maximization .....	39.8	-19.0	2
D—Status maximization .....	43.8	-23.0	2
E—Status maximization .....	44.4	-13.2	3
Country-specific models:			
A1—Baseline: Country-specific weight .....	2.8	28.4	3
B1—Acculturation: Country-specific weight . . . .	40.4	1.2	4
B2—Acculturation: Country-specific acculturation .....	46.7	-5.1	4
B3—Acculturation: Country-specific weight and acculturation .....	48.5	24.4	7
Class-specific models			
A2—Baseline: Class-specific weight .....	11.4	40.7	5
A3—Baseline: Class- and country-specific weight	47.1	192.4	23
B4—Acculturation: Class-specific weight .....	52.8	9.7	6
B5—Acculturation: Class-specific acculturation...	49.6	12.9	6
B6—Acculturation: Class-specific weight and acculturation .....	59.3	55.2	11

NOTE.— $N = 33,222$ . Results controlled for respondent's religious denomination and age and for the year of the survey. Equations for the models are given in appendix A. The  $G^2$ , BIC, and  $df$  used are reported for the baseline model, for the nested models the reduction in  $G^2$  (against the  $df$ ) and BIC compared to the baseline model are reported.

The first model to be compared with this baseline model is the one including an acculturation effect. The acculturation hypothesis predicts that the older one is, the larger will be the influence of the destination class relative to that of the origin class. In other words, this hypothesis implies that the weight parameters vary according to the length of time that one is a member of a class. Model B therefore includes an interaction between age and the origin and destination weights. The equation of this model is

$$\text{prob}(Y_{ijk} = 1) = 1/(1 + e^{-lin}); \quad (1)$$

$$\begin{aligned} lin = & (p + \delta p \times \text{AGE})\alpha_i \\ & + [(1 - p) - \delta p \times \text{AGE}]\alpha_j + \beta_L \text{cov}_L. \end{aligned} \quad (4)$$

According to this model the destination weight increases by  $\delta p$  for each extra year of age. Conversely, the origin weight decreases by  $\delta p$  for each extra year of age. This symmetry means that  $[p + (1 - p)]$  still sums to 1 (see De Graaf 1991).

On the standard criteria, this acculturation model results in a significant improvement compared to our baseline model. Given our very large sample sizes, it can be questioned whether the conventional criteria are altogether appropriate, and we therefore report the BIC statistic too (Raftery 1986), which adjusts the fit statistic by sample size and degrees of freedom. The more negative the BIC, the better the fit of the model. Table 3 shows that both on the  $G^2$  statistic ( $-\Delta G^2/\Delta df$ ) and the BIC criterion, the acculturation hypothesis (model B) is to be preferred to the baseline model.<sup>15</sup>

Next, the status maximization hypothesis predicts that for the downwardly mobile the process of acculturation will take longer than it does for the upwardly mobile. To test this, we have to define a variable ( $UP$ ) that indicates whether one is upwardly mobile or downwardly mobile. Determining the direction of mobility is not straightforward, given the categorical class scheme. We assume that those who move from any other class to the service class are upwardly mobile (i.e., those with origin class 2, 3, 4, 5, or 6 and destination class 1 have a score of "1" and others have a score of "0").

The next question is how to test this hypothesis, since there are three

<sup>15</sup> Since the diagonal model can also be considered as a restricted version of Duncan's (1966) square additive model, we also tested the logistic version of the square additive model. This model resulted in a BIC of  $-305,886$ , while using 16  $df$  more than the diagonal model. The baseline diagonal model resulted in a BIC of  $-306,023$ . Therefore, besides theoretical reasons, we also prefer the diagonal version instead of the conventional model for statistical reasons.

possible interpretations (see Nieuwbeerta and De Graaf 1993). One interpretation is that the rate of adjustment is the same among the upwardly and downwardly mobile and that it is the extent of adjustment that differs. Given this interpretation we can model an additional interaction effect on the weight parameters. This is model C in table 3.<sup>16</sup> A second interpretation is that the rate of acculturation is simply slower for the downwardly mobile. Because the upwardly mobile gain status, they will adapt faster to their class of destination than will the downwardly mobile. We can test this interpretation by adding a second-order interaction term between the variables UP and AGE. This is done in model D in table 3. A third interpretation is that the first two interpretations are simultaneously valid. This interpretation is expressed in model E in table 3.

As the BIC statistic shows, none of these three interpretations is preferable to the acculturation model B. We should, therefore, reject the hypothesis that the downwardly mobile and upwardly mobile differ in their acculturation pattern.<sup>17</sup> This result contradicts earlier empirical studies (for an overview see Janowitz [1970]) but is in line with the more recent empirical results of De Graaf and Ultee (1990), Weakliem (1992), and Nieuwbeerta and De Graaf (1993).

The conclusion so far is that the acculturation model B gives the best representation of the data. We should note that similar nested model comparisons for each country separately also showed the acculturation model to be preferable. The acculturation model B assumes that there are no differences between our four countries in the strength of the destination and acculturation effects. This assumption, however, may not be warranted, and there may be country-specific mobility effects in the same way that there may be some country-specific effects of class, religion, or period. Our next step, therefore, is to test whether there are differences in these effects between countries. The first country-specific model that

<sup>16</sup> Given our definition of upward and downward mobility, we have a large number of respondents who are mobile but neither upwardly nor downwardly. In models C, D, and E we included an interaction term that only models an extra destination effect for the upwardly mobile. A more strict test of our hypothesis would be to include an extra interaction term for the downwardly mobile as well. This would imply that the reference category comprises those who are mobile but neither upward nor downward. However, these more complex models did not result in any significant improvement.

<sup>17</sup> We would like to note that the statement "We should, therefore, reject the hypothesis that . . ." is not a standard statement, because the hypothesis being rejected is an alternative hypothesis. Standard hypothesis testing allows one only to reject a null hypothesis; in this framework one cannot reject an alternative hypothesis, but can only fail to reject the null. In the Bayesian approach that underlies BIC, however, it is perfectly valid to say that the data reject an alternative hypothesis, since they provide evidence for the null. We are very grateful to an *AJS* reviewer for this observation.

we test assumes that the weight parameter varies among the countries (model A1 in table 3). This model does not lead to a significant improvement, either on the  $G^2$  or the BIC criterion, over model A. We next turn to variations on the acculturation model (i.e., models B1, B2, and B3). These models fit different weight parameters or different acculturation parameters for each country. As table 3 shows, none of these models results in significant improvement in fit compared with model B.<sup>18</sup> Hence, our conclusion is that both destination/origin effects and acculturation effects are similar in all four countries under investigation.

The parameter estimates of our preferred model (acculturation model B) are presented in table 4. The estimates for the covariates indicate that religious denomination is an important predictor in all four countries. In particular the Orthodox Reformed in the Netherlands have a relatively low probability of having a left-wing political preference. Age has no significant impact in Germany, the Netherlands, and the United States. Not unexpectedly, in the United States blacks have a higher chance of having a left-wing preference than nonblacks.

The estimates of the diagonal reference parameters are presented in table 4 as well. In order to help the reader interpret the parameter estimates, we also calculate the predicted probability of having a left-wing preference within each reference class for respondents who are 18–25 years old, nonblack, and nonreligious in the reference period. The diagonal parameters show that, with the exception of the United States, the intergenerationally stable petty bourgeois and the farmers have the lowest probability of having a left-wing preference. In all four countries the manual workers and the agricultural workers have the highest probability of having a left-wing preference.

Our main interest, however, is in the weight parameter  $p$  and the acculturation parameter  $\delta p$ . The weight parameter is .585 (SE = .018) and the acculturation parameter is .005 (SE = .001). This suggests that the cumulative impact of acculturation over the life span is indeed substantial. Thus our preferred model suggests that for 18 year olds (coded –22 on our AGE variable) in all four countries the relative destination weight is  $.585 - 22 \times .005$ , that is .475, and the origin weight is .525. This implies that for our youngest respondents the effect of origin is slightly more important than the effect of their current class. By the time

<sup>18</sup> We experimented with various combinations of country-specific destination weights and country-specific acculturation effects. Since our analyses for each country separately showed, especially for the United States, a relatively strong acculturation effect, we tested, for example, for U.S. exceptionalism. The upshot of all these nested model comparisons is that the simple acculturation model B fits the data best, both by the  $G^2$  and BIC standards.

TABLE 4  
PARAMETER ESTIMATES OF THE ACCULTURATION MODEL B

	Germany	Great Britain	Netherlands	United States
$\alpha$ : Political Preference of Immobile				
Service class .....	-.753*** (32)	-2.017*** (12)	-.655*** (34)	-.544** (20)
Routine nonmanual .....	-.416* (40)	-1.139*** (24)	-.054 (49)	-.353 (24)
Petty bourgeois .....	-2.087*** (11)	-2.308*** (9)	-1.193*** (24)	-.323 (24)
Farmers.....	-2.325*** (8)	-2.141*** (10)	-1.620*** (17)	-.255 (33)
Working class .....	.456** (61)	.339* (58)	.987*** (73)	.181 (37)
Agricultural laborers ...	-.074 (48)	-.555 (37)	.442 (61)	.370 (38)
$\beta$ : Effect of Covariate on Political Preference				
Religion:				
Catholic.....	-.953***	.444*	-1.513***	-.217
Reformed/Protestant.	-.227	-.260*	-.783***	-.656***
Orthodox Reformed	..	...	-2.749***	...
Other .....	-.337	.609*	-.632*	.382
No religion. ....	0	0	0	0
Age:				
18-30.....	.135	-.123	.044	-.151
31-50 <sup>a</sup> .....	0	0	0	0
51+ .....	-.062	-.293*	.259	.130
Ethnicity:				
Nonblack <sup>a</sup> .....	.	..	...	0
Black.....	..	...	..	1.697***
Period:				
1961-65.....	...	.691***	..	...
1966-70.....	1.023***	.849***	-.248	...
1971-75.....	.614***	.524**	.305	-.164
1976-80.....	.575***	.175	.143	.490***
1981-85 <sup>a</sup> .....	0	0	0	0
1986-90.....	.408**	-.082	-.417***	-.309**

NOTE — Destination weight ( $\phi$ ) equals .585\*\*\*, effect of age on destination weight ( $\delta\phi$ ) equals .005\*. Numbers in parentheses report probability of having a left-wing preference within the reference category.

<sup>a</sup> Reference category

\*  $P < .05$

\*\*  $P < .01$

\*\*\*  $P < .001$ .

they have reached 65 years of age (coded 25 on our AGE variable), the relative destination effect becomes  $.585 + 25 \times .005$ , that is .710, and the origin effect becomes .290. Although these figures indicate that the destination effect is over twice the size of the origin effect, origin still has some impact on the political preferences of 65-year-olds.

## RESULTS: MACRO EFFECTS OF INFLOW MOBILITY

### Compositional Effect

Table 2 confirms that a compositional effect of social mobility is present and suggests that intergenerational mobility diminishes the "democratic class struggle." For example, in Britain 61.2% of the stable members of the manual working class had a left-wing political preference, whereas the figure for the class as a whole falls to 56.4%. Hence, the compositional effect amounts to  $-4.8\%$ . Conversely, the proportion of the service class as a whole with a left-wing preference (19.7%) is greater than that in the core service class (13.1%). In this case the compositional effect is  $6.6\%$ . In general, then, we find that the class differences down the main diagonal (that is, the differences between the core classes) are greater than the differences between the classes taken as a whole. This is the case for all three European countries, although it is less evident in the case of America. (Since class has a weaker effect on voting behavior in the United States, the compositional effects of social mobility, other things being equal, are likely to be smaller too.)

Clearly, then, there are compositional effects, but as we argued earlier the size of these effects will depend not only on the amount of inflow but also on the political complexion of that inflow. For example, in the case of Britain 24% of the manual class were newcomers coming from other class origins, but the proportion of these newcomers with left-wing political preferences was only .41, rather different from the proportion (.61) of stable members who had left-wing political preferences. In the service class, on the other hand, 68% were newcomers, but their political preferences were not so deviant from those of the stable members: 23% of newcomers to the service class had a left-wing political preference, compared with 13% of the stable members. The overall compositional effect on the service class is thus  $(23\% - 13\%) \times 68\%/100 = 7\%$  (which is of course the same as the difference between the diagonal figure and the figure for all service-class workers in table 2), while the overall compositional effect in the manual class is  $(41\% - 61\%) \times 24\%/100 = -5\%$ . The greater volume of inflow into the service class, then, is partly canceled out by the lesser political deviance of that inflow, and the net result is that the compositional effect on the service class is not notably large,

despite the heterogeneity in its social origins. Similar results can be obtained for the other European countries. Indeed, in Germany the compositional effect in the service class (5%) is actually smaller than that in the manual working class (−6%) despite the much greater volume of mobility into the service class.

This compositional account ignores any processes of mutual interaction and influence between the stable and mobile members of a class. It simply takes the political preferences of the mobile and immobile as given. We now move on to consider whether we can find any evidence for contextual processes that might help to account for the political preferences of the mobile and immobile respectively.

### Contextual Effect for the Mobile

Our first contextual hypothesis states that a class with a low level of inflow mobility (and therefore a high level of demographic identity) will have a greater impact on newcomers than will a class with a higher level of inflow mobility. First, however, we must investigate whether the social classes vary in their impact on newcomers at all before we move on to test whether the differing impacts on newcomers can be explained by the rates of inflow mobility.

To calculate the impact of a class on newcomers, we fit models A2 and A3 (see table 3). Model A2 fits a separate weight parameter for each destination class. The class-specific weight parameters that we obtain from this model thus give us estimates of the different impacts of the classes on newcomers. However, as we can see from table 3, this model does not yield a significant improvement in fit over the baseline model A. Model A3, which allows each country to have different class-specific weights, is no more successful.

We have also experimented with various acculturation models that fit class-specific effects (models B4–B6 in table 3). For example, in model B5 we test whether it is the acculturation process (that is, the interaction with age) that varies between classes. None of our investigations,<sup>19</sup> however, yielded a class-specific version that unambiguously displayed improved fit compared with the corresponding class-uniform model. The BIC statistic clearly suggests that model B is the best, while the standard hypothesis tests indicate that B4, with class-specific weight parameters, and B6, with class-specific weight and acculturation parameters, do fit significantly better (B – B4:  $G^2 = 13.0$ ,  $df = 5$ ; B – B6:  $G^2 = 19.5$ ,

<sup>19</sup> For the sake of simplification, we present only a few of the models we tested.

$df = 10$ ). For the moment we should perhaps use the Scottish judicial verdict of "not proven." At any rate, we do not yet have sufficiently good grounds to proceed with tests of the contextual hypothesis for the mobile; if there are no class differences in impacts on newcomers, there is nothing for us to explain, and we must a fortiori reject our first contextual hypothesis for the mobile.

#### Contextual Effect for the Immobile

Our second contextual hypothesis implies that the extent and political character of the inflow mobility will influence the political preferences of the stable class members. Here, therefore, our interest is in explaining the variations in the political preferences of the intergenerationally stable members of each class. There are, as we saw in table 2, large differences in the political preferences of these core classes, but we must of course recognize that these are likely to be due to the distinct material circumstances and interests of the classes and not solely to mobility patterns. We take this into account in our modeling by allowing each class within each nation to have its own "natural" level of left-wing preferences. (This is why we also estimated country-specific diagonal cells for our models in table 3). In this way we can test whether rates of inflow mobility can account for variations around this "natural" level.

The essence of our model, then, is that the level of left-wing preference in a particular year among the stable members of a given class in a country is to be explained by the level of inflow into that class in that year. In order to take account of the political complexion of the inflow, we have to weight the inflow. The basic ideas can be represented by:

$$\pi_{yin} = \text{PREF}_{in} + \gamma (\text{INFLOW}_{yin} \times \text{WEIGHT}_{in}) + E_{yin}, \quad (5)$$

where  $\pi_{yin}$  is the proportion of the stable members in class  $i$ , year  $y$ , and nation  $n$  who support a left-wing party;  $\text{PREF}_{in}$  gives the "natural" level of left-wing political preferences among the stable members of class  $i$  in nation  $n$  (and is measured by  $[6 \times 4 - 1]$  23 dummy variables);  $\text{INFLOW}_{yin}$  represents the proportion of class  $i$  in nation  $n$  and year  $y$  who were intergenerationally mobile (that is, who came from social origins other than  $i$ ). There are many possible procedures that could be used for calculating  $\text{WEIGHT}_{in}$ , but the one reported here uses the political preferences of the stable members of the origin classes.<sup>20</sup>

<sup>20</sup> A numerical example may clarify our procedure. In calculating the left-wing inflow mobility into the service class in Britain in 1964, we first take the actual proportions of the service class in Britain in 1964 who were mobile from the other five class origins. From the manual class, 47.2% were mobile. We then weight this proportion by .612 (the proportion of the stable members of the working class who had a left-wing political preference in the combined British data set, as shown in table 2) From the



In this model the unit of analysis is not the individual but the class (within a given nation in a given year). We therefore have 282 observations (and we naturally must weight these observations according to the number of individuals on which they are based). Our dependent variable is the proportion of that class with left-wing political preferences and is therefore a continuous variable (and is approximately normally distributed). We can therefore use linear regression. The model explains 29% of the variance, but the key parameter of interest, the parameter  $\gamma$ , is nowhere near significant ( $t = 1.1$ ). We therefore have to reject the hypothesis that the more left-wing mobility there is into a class, the more likely are the immobile members to have a left-wing political preference.

Unlike our earlier microlevel analyses, the macrolevel model represented in equation (4) fails to control for the other characteristics of the stable class members, such as their religion, age, and ethnicity. However, further modeling that incorporates controls for these characteristics reaches essentially the same conclusions.<sup>21</sup>

---

agricultural worker class, 4.1% were mobile, and we weighted this proportion by .409. The weighted inflow to the service class in Britain in 1964 thus becomes  $4.1\% \times .409 + 47.2\% \times .612 + 4.2\% \times .079 + 11.3\% \times .113 + 7.4\% \times .326 = 34.584\%$ . These weights must be kept constant over time; if we allow them to vary each year, we introduce a potential risk of circularity into the model: In some years, quite independent of levels of mobility, there may be across-the-board increases in left-wing political preferences affecting the mobile and the stable alike. Weights that were allowed to vary each year might thus be correlated with the yearly variations in the dependent variable. We are grateful to an *AJS* reviewer for suggesting this method of constructing our explanatory variable.

<sup>21</sup> One way to control for these individual-level characteristics is to use a multilevel model (Burstein, Linn, and Capell 1978; Goldstein 1987). The first level becomes an individual-level model in which we use exactly the same control variables as we did earlier in the article. Our contextual model, expressed in eq. (4), then becomes the second level. Within the multilevel framework, the dependent variable in this second level is no longer the observed proportion with a left-wing preference but is the intercept (for the given level-2 unit) from the level-1 equation. This intercept can be thought of as the predicted probability of supporting a left-wing political party in that particular level-2 unit for someone with the baseline individual characteristics. More formally, the level-1 equation is

$$Y_{iwy} = \zeta_{iwy} + \eta_{Lw} \text{cov}_{Lw} + E_{iwy}, \quad (6)$$

where  $\zeta_{iwy}$  represents the intercepts for the level-2 units and there are  $L$  covariates, as in eq. (2). The level-2 equation is

$$\zeta_{iwy} = \text{PREF}_{iw} + \theta (\text{INFLOW}_{iwy} \times \text{WEIGHT}_{iw}) + u_{iwy} \quad (7)$$

This is what is known as a random intercepts model. As with microlevel analysis, it is appropriate to employ a logistic reformulation of these equations, and this can be done easily. The analyses were conducted using the VARCL computer program. Within VARCL, eqq. (5) and (6) are estimated simultaneously on the basis of the

## CONCLUSIONS AND DISCUSSION

In this article we modeled the impact of intergenerational class mobility on political preference. For this purpose we used 55 cross-sectional data sets representing the United States, Britain, Germany, and the Netherlands. These data cover the period 1964–90. We distinguished hypotheses on the micro as well as the macro level. On the micro level we tested three hypotheses using the design of Sobel's diagonal reference models. We found that an acculturation model resulted in the best fit. Furthermore, a test of possible differences between the four countries revealed that an acculturation model with the same destination weight and acculturation process in all four countries gave the best fit.

This commonality in the relative effects of mobility is a surprising and important result in its own right. We might a priori expect to find substantial differences between the more traditional societies of Europe, where class distinctions are believed to be more deeply ingrained, than in a new society such as the United States. In the United States we might have anticipated that the relative effects of origin might be smaller and the acculturation process more rapid. But despite our large data set, we could not reject the hypothesis that the same model applied to all four of our countries.

To be sure, there are large absolute differences between our four countries in the size of the class effects on voting, and our results are the same as those of other scholars who have shown that class voting is weaker in the United States. The important finding is that the *relative* importance of origin and destination for political preferences is not significantly different in America. This result mirrors findings about the mobility process itself: While absolute rates of mobility tend to be higher in the United States, relative mobility rates, or fluidity (as measured by the odds ratios), is not exceptionally high in the United States (Erikson and Goldthorpe 1985).

On the macro level we tested one compositional and two contextual hypotheses. Our results did indeed show the force of compositional arguments. In general, mobility tends to weaken the "democratic class struggle," but the effects of the amount of inflow mobility depend on the political complexion of that inflow. While the service class, for example, is a great deal more heterogeneous in its social origins than is the manual working class in all four of our countries, we found that in the European ones this heterogeneity was largely canceled out by the relatively right-wing character of the inflow into the service class.

---

maximum-likelihood procedure (for a more detailed account, see Longford [1988]). However, the results of the multilevel logistic analysis are essentially the same as those of the simpler analysis reported in the text, and the parameter for the weighted inflow,  $\theta$ , remains nonsignificant ( $t = 0.43$ ).

While the occurrence of compositional effects is not, perhaps, very surprising, it is more interesting to observe that they fail to account for the relatively low class polarization in the United States. Even when we look at the stable members of each class in the United States (the diagonal cells), we find that class differences are small.

While the compositional effects of social mobility can clearly be quite substantial, however, our article found little evidence for the more subtle ways in which mobility can reduce class polarization through contextual processes. We were not able to provide convincing support for contextual effects either on the mobile or on the immobile. One possible explanation for this is that contextual class processes simply are weak in contemporary industrial societies. Contextual processes of the kind we envisaged may require stable communities with fairly clear boundaries, high levels of in-group interaction, and low levels of out-group interaction. These conditions may be met for some social groups in contemporary society but may be weak or absent for social classes. For example, there does seem to be some evidence for contextual school effects, but it is likely that schoolchildren interact much more with pupils from the same school than with pupils from other schools. The social conditions under which contextual effects emerge may, therefore, not be satisfied in the case of social classes.

Alternatively, it may be that we have focused on the wrong aspects of class. For example, patterns of intragenerational mobility and stability may be more important than those of intergenerational mobility. Intergenerational mobility tables give us only a snapshot; some people who, in the snapshot, are intergenerationally stable may subsequently move. This is particularly likely in a class like the routine nonmanual workers, membership of which tends to be rather transitory. Conversely, intergenerationally mobile newcomers to the service class may have high probabilities of intragenerational stability once they have arrived. It may therefore be that the core of the class, which defines its norms and values, will consist of those who are indeed long-term members, but intergenerational stability may not be a good indicator of long-term membership.

Another possibility is that the processes we hypothesized do occur but over a much longer time scale. The more diffuse the patterns of social interaction are, the longer it may take for mobility rates to shape class processes. Processes that may emerge in a matter of months under the conditions of intensive within-group interaction that may characterize schools may take a generation to emerge under the conditions that characterize social classes. Our evidence suggests that the process of individual acculturation to the political preferences of the destination class is a gradual and long-term process, continuing over the life span. The same may well be true of the contextual processes.

It is too soon, then, to reject the contextual theories of class formation and the expressive theories of voting behavior with which we associated them. After all, the existence of acculturation is difficult to explain except by the effects of social context. Ideally, too, we would use direct measures of people's patterns of social interaction in addition to the indirect measures based on mobility rates used in this article. What we hope to have shown is that the older tradition of macrosociological thinking about the effects of mobility raises questions that are both researchable and are of continuing interest.

## APPENDIX A

### Definitions of Models

#### *General Models*

$$A: \text{lin} = (p)\alpha_{in} + (1-p)\alpha_{jn} + \beta_L \text{cov}_L$$

$$B: \text{lin} = (p + \delta p \times \text{AGE})\alpha_{in} \\ + [(1-p) - \delta p \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

$$C: \text{lin} = (p + \delta p1 \times \text{AGE} + \delta p2 \times \text{UP})\alpha_{in} \\ + [(1-p) - \delta p1 \times \text{AGE} - \delta p2 \times \text{UP}]\alpha_{jn} + \beta_L \text{cov}_L$$

$$D: \text{lin} = (p + \delta p1 \times \text{AGE} + \delta p2 \times \text{UP} \times \text{AGE})\alpha_{in} \\ + [(1-p) - \delta p1 \times \text{AGE} - \delta p2 \times \text{UP} \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

$$E: \text{lin} = (p + \delta p1 \times \text{AGE} + \delta p2 \times \text{UP} + \delta p3 \times \text{UP} \times \text{AGE})\alpha_{in} \\ + [(1-p) - \delta p1 \times \text{AGE} - \delta p2 \times \text{UP} - \delta p3 \times \text{UP} \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

#### *Nation-Specific Models*

$$A1: \text{lin} = p_n \alpha_{in} + (1-p_n)\alpha_{jn} + \beta_L \text{cov}_L$$

$$B1: \text{lin} = (p_n + \delta p \times \text{AGE})\alpha_{in} \\ + [(1-p_n) - \delta p \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

$$B2: \text{lin} = (p + \delta p_n \times \text{AGE})\alpha_{in} \\ + [(1-p) - \delta p_n \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

$$B3: \text{lin} = (p_n + \delta p_n \times \text{AGE})\alpha_{in} \\ + [(1-p_n) - \delta p_n \times \text{AGE}]\alpha_{jn} + \beta_L \text{cov}_L$$

*Class-Specific Models*

$$A2: \text{lin} = p_i \alpha_{in} + (1 - p_i) \alpha_{jn} + \beta_L \text{cov}_L$$

$$A3: \text{lin} = p_{in} \alpha_{in} + (1 - p_{in}) \alpha_{jn} + \beta_L \text{cov}_L$$

$$B4: \text{lin} = (p_i + \delta p \times \text{AGE}) \alpha_{in} \\ + [(1 - p_i) - \delta p \times \text{AGE}] \alpha_{jn} + \beta_L \text{cov}_L$$

$$B5: \text{lin} = (p + \delta p_i \times \text{AGE}) \alpha_{in} \\ + [(1 - p) - \delta p_i \times \text{AGE}] \alpha_{jn} + \beta_L \text{cov}_L$$

$$B6: \text{lin} = (p_i + \delta p_i \times \text{AGE}) \alpha_{in} \\ + [(1 - p_i) - \delta p_i \times \text{AGE}] \alpha_{jn} + \beta_L \text{cov}_L$$

DATA REFERENCES

- Allerbeck, K. R., M. Kaase, H. D. Klingemann, F. J. Heunks, P. C. Stouthard, J. J. A. Thomassen, J. W. von Deth, S. H. Barnes, B. G. Farah, R. Inglehart, and M. K. Jennings. 1980. *Political Action II*. Distributed by the Zentralarchiv fuer Empirische Sozialforschung, University of Cologne, Cologne, ZA 1188.
- Arts, K., E. Hollander, K. Renckstorf, and P. Verschuren. 1989. *Grootschalig onderzoek naar media-uitrusting, media-exposure en mediagebruik in Nederland, 1989*. Data obtained from the authors.
- Barnes, S. H., and M. Kaase. 1976. *Political Action: An Eight Nation Study, 1973-1976*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7777.
- Butler, D., and D. E. Stokes. 1963-66. *Political Change in Britain, 1963-1970: (1) Cross-Section Sample, 1963; (2) Electorate Sample, 1964; (3) Electorate Sample, 1966*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7232-7250-7233-7234.
- . 1970. *Political Change in Britain, 1969-1970*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7004.
- CBS (Centraal Bureau voor de Statistiek). 1977. *Life (Situation) Survey, Netherlands, 1977*. Distributed by the Steinmetz Archive, Amsterdam. P0328.
- Crewe, I., B. Saerlvik, and J. Alt. 1974. *British Election Study: 1974, October*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7870.
- Crewe, I., B. Saerlvik, and D. Robertson. 1979. *British Election Study: 1979 Cross-Section, May*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 8196.
- Davis, J. A., and T. W. Smith. 1991. *General Social Survey Cumulative File, 1972-1991*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 9505.
- Eisinga, R., A. Felling, J. Peters, P. Scheepers, and O. Schreuder. 1990. *Religion in Dutch Society 90*. Distributed by the Steinmetz Archive, Amsterdam. P1100.
- Felling, A., J. Peters, and O. Schreuder. 1985. *Religion in Dutch Society 85*. Distributed by the Steinmetz Archive, Amsterdam. P1012.
- Heath, A. F., R. Jowell, J. Curtice, and J. Field. 1983. *British Election Study, (June)*

1983. Distributed by the Social Science Research Council. University of Essex, Essex. 2005.
- Heath, A. F., R. Jowell, J. Curtice, J. Field, and S. Witherspoon. 1989. *British Election Study, 1987*. Distributed by the Social Science Research Council. University of Essex, Essex. SN 2568.
- Hermkens, P., and P. Van Wijngaarden. 1976. *Criteria for Justification of Income Differences, Netherlands, 1976*. Distributed by the Steinmetz Archive, Amsterdam. P0653.
- Heunks, F. J., M. K. Jennings, W. E. Miller, P. C. Stouthard, and J. Thomassen. 1973. *Dutch Election Study, 1970–1973*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7261.
- International Social Science Program. 1987. *Social Inequality, 1987*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 9383.
- Kaase, M., and U. Schleth. 1969. *German Election Study: 1969, August–September*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7108. Also distributed by the Zentralarchiv fuer Empirische Sozialforschung, University of Cologne, Cologne. ZA 0525.
- Klingemann, H. D., and F. U. Pappi. 1969. *German Pre- and Post-election Study, 1969*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7098.
- Van der Eijk, C. B., G. A. Irwin, and B. Niemoeller. 1986. *Dutch Parliamentary Election Study, 1986*. Distributed by the Steinmetz Archive, Amsterdam. P0866A.
- Van der Eijk, C. B., B. Niemoeller, and A. T. J. Eggen. 1981. *Dutch Parliamentary Election Study, 1981*. Distributed by the Steinmetz Archive, Amsterdam. P0350.
- Van der Eijk, C. B., B. Niemoeller, M. J. Koopman, and S. M. J. Koopman. 1982. *Dutch Parliamentary Election Study, 1982*. Distributed by the Steinmetz Archive, Amsterdam. P0633A.
- Verba, S., N. H. Nie, and J. Kim. 1971. *Cross-National Program in Political and Social Change: Netherlands Cross-Sectional Survey*. Distributed by the Inter-university Consortium for Political and Social Research, Ann Arbor, Mich. ICPSR 7768.
- Werkgroep Nationaal Kiezersonderzoek. 1977. *Dutch Parliamentary Election Study, 1977*. Distributed by the Steinmetz Archive, Amsterdam. P0354.
- ZUMA (Zentrum fuer Umfragen, Methoden, und Analysen). 1980. *ZUMA-Standarddemographie (Zeitreihe), Germany, 1976–1980*. Distributed by the Zentralarchiv fuer Empirische Sozialforschung, University of Cologne, Cologne. ZA 1233.
- . 1988. *German Social Survey (Allbus) Cumulative File, 1980–1988*. Distributed by the Zentralarchiv fuer Empirische Sozialforschung, University of Cologne, Cologne. ZA 1795.
- . 1990. *German Social Survey (Allbus), 1990*. Distributed by the Zentralarchiv fuer Empirische Sozialforschung, University of Cologne, Cologne. ZA 1800.

## REFERENCES

- Abramson, P. R. 1972. "Intergenerational Social Mobility and Partisan Choice." *American Political Science Review* 66:1291–94.
- Barber, J. A., Jr. 1970. *Social Mobility and Voting Behavior*. Chicago: University of Chicago Press.
- Bartolini, S., and P. Mair. 1990. *Identity, Competition and Electoral Availability: The Stabilization of European Electorates, 1885–1985*. Cambridge: Cambridge University Press.

- Blau, P. M. 1956. "Social Mobility and Interpersonal Relations." *American Sociological Review* 21:290-95.
- . 1964. *Exchange and Power in Social Life*. New York: Wiley.
- Blau, P. M., and J. E. Schwartz. 1984. *Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations*. Orlando, Fla.: Academic Press.
- Burstein, L., R. L. Linn, and F. J. Capell. 1978. "Analyzing Multilevel Data in the Presence of Heterogeneous Within-Class Regressions." *Journal of Educational Statistics* 3:347-83.
- Clark, T. N., and S. M. Lipset. 1991. "Are Social Classes Dying?" *International Sociology* 6:397-410.
- Clifford, P., and A. Heath. 1993. "The Political Consequences of Social Mobility." *Journal of the Royal Statistical Society* 156:1-11.
- Converse, P. E. 1958. "The Shifting Role of Class in Political Attitudes and Behavior." Pp. 388-99 in *Readings in Social Psychology*, 3d ed. Edited by E. E. Mac-coby, T. M. Newcomb, and E. L. Hartley. New York: Holt.
- De Graaf, N. D. 1991. "Distinction by Consumption in Czechoslovakia, Hungary, and the Netherlands." *European Sociological Review* 7:267-90.
- De Graaf, N. D., and H. B. G. Ganzeboom. 1990. "Cultuurdeelname en opleiding: een analyse van status groep-effecten met diagonale referentiemodellen" (Culture consumption and education: An analysis of status-group effects with diagonal reference models). *Mens en Maatschappij* 65:2-27.
- De Graaf, N. D., and A. F. Heath. 1992. "Husbands' and Wives' Voting Behavior in Britain. Class-Dependent Mutual Influence of Spouses." *Acta Sociologica* 35:311-22.
- De Graaf, N. D., and W. Ultee. 1990. "Individual Preferences, Social Mobility and Electoral Outcomes." *Electoral Studies* 9:109-32.
- Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Duncan, O. D. 1966. "Methodological Issues in the Analysis of Social Mobility." Pp. 51-97 in *Social Structure and Mobility in Economic Development*, edited by N. L. Smelser and S. M. Lipset. Chicago: Aldine.
- Durkheim, É. 1897. *Le Suicide: Étude de sociologie*. Paris: Alcan.
- Erikson, R., and J. H. Goldthorpe. 1985. "Are American Rates of Social Mobility Exceptionally High? New Evidence on an Old Issue." *European Sociological Review* 1:1-22.
- . 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon.
- Erikson, R., J. Goldthorpe, and J. Portocarero. 1979. "Intergenerational Class Mobility in Three Western European Societies." *British Journal of Sociology* 30:415-41.
- Franklin, M. N., T. Mackie, and H. Valen, eds. 1992. *Electoral Change: Responses to Evolving Social and Attitudinal Structures in Western Countries*. Cambridge: Cambridge University Press.
- Ganzeboom, H. B. G., R. Luijkx, and D. J. Treiman. 1989. "Intergenerational Class Mobility in Comparative Perspective." *Research in Social Stratification and Mobility* 8:3-84.
- Girod, R. 1971. *Mobilité sociale: Faits établis et problèmes ouverts*. Geneva: Droz.
- Goldstein, H. 1987. *Multilevel Models in Educational and Social Research*. London: Griffin.
- Goldthorpe, J. H. 1980. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon.
- Heath, A. F., R. Jowell, and J. Curtice. 1985. *How Britain Votes*. Oxford: Pergamon.
- Heath, A. F., R. Jowell, J. Curtice, G. Evans, J. Field, and S. Witherspoon. 1991. *Understanding Political Change*. Oxford: Pergamon.
- Heath, A. F., B. Taylor, and G. Toka. 1993. "Religion, Morality and Politics." Pp. 49-80 in *International Social Attitudes: The 10th ISA Report*, edited by R. Jowell, L. Brook, and L. Dowds. Aldershot: Dartmouth.

- Hendrickx, J., N. D. De Graaf, J. Lammers, and W. Ultee. 1993. "Models for Status Inconsistency and Mobility: A Comparison of the Approaches by Hope and Sobel with the Mainstream Square Additive Model." *Quality and Quantity* 27:335-52.
- ILO (International Labour Office). 1969. *International Standard Classification of Occupations*, rev. ed., 1968. Geneva: International Labour Office.
- Janowitz, M. 1970. *Political Conflict*. Chicago: Quadrangle.
- Kelley, J. 1992. "Social Mobility and Politics in the Anglo-American Democracies." Pp. 21-50 in *Social Mobility and Political Attitudes: Comparative Perspectives*, edited by F. Turner. New Brunswick, N.J.: Transaction.
- Knoke, P. 1973. "Intergenerational Occupational Mobility and the Political Party Preferences of American Men." *American Journal of Sociology* 78:1448-68.
- Lipset, S. M. 1960. *Political Man: The Social Bases of Politics*. London: Heinemann Educational Books.
- Lipset, S. M., and R. Bendix. 1959. *Social Mobility in Industrial Society*. Berkeley: University of California Press.
- Lipset, S. M., P. F. Lazarsfeld, A. H. Barton, and J. Linz. 1954. "The Psychology of Voting: An Analysis of Political Behavior." Pp. 1124-75 in *Handbook of Social Psychology*, edited by G. Lindzey. Cambridge, Mass.: Addison-Wesley.
- Lipset, S. M., and H. Zetterberg. 1956. "A Theory of Social Mobility." *Transactions of the Third World Congress of Sociology* 3:155-77.
- Longford, N. T. 1988. *VARCL: Software for Variance Component Analysis of Data with Hierarchically Nested Random Effects*. Princeton, N.J.: Educational Testing Service.
- Marx, K. 1926. *The Eighteenth Brumaire of Louis Bonaparte*, translated by P. Eden and P. Cider. London: Allen Unwin.
- Nieuwbeerta, P., and N. D. De Graaf. 1993. "Intergenerational Class Mobility and Political Preference in the Netherlands between 1970 and 1986." *Netherlands Journal of Social Sciences* 29:27-45.
- Nieuwbeerta, P. 1995. *The Democratic Class Struggle in 20 Countries, 1945-1990*. Ph.D. dissertation. University of Nijmegen. In press.
- Parkin, F. 1971. *Class, Inequality and Political Order: Social Stratification in Capitalist and Communist Societies*. London: Palladin.
- Raftery, A. 1986. "Choosing Models for Cross-Classifications: Comment on Grusky and Hauser." *American Sociological Review* 51:145-46.
- Sobel, M. E. 1981. "Diagonal Mobility Models: A Substantively Motivated Class of Designs for the Analysis of Mobility Effects." *American Sociological Review* 46:893-906.
- . 1985. "Social Mobility and Fertility Revisited. Some New Models for the Analysis of the Mobility Effects Hypothesis." *American Sociological Review* 50 699-712.
- Sorokin, P. A. 1957. *Social and Cultural Mobility*. Glencoe, Ill.: Free Press.
- Thorburn, P. 1979. "Class and Mobility: The Political Preferences of Men in England and Wales." Ph.D. dissertation. University of Michigan, Department of Political Science.
- Turner, F. 1992. *Social Mobility and Political Attitudes: Comparative Perspectives*. New Brunswick, N.J.: Transaction.
- Ultee, W. C., and N. D. De Graaf. 1991. "Cultuurdeelname en opleidingsverschillen in primaire relaties. De invloed van intergenerationele mobiliteit en gemengde huwelijken" (Cultural consumption in primary relations: The influence of intergenerational mobility and mixed marriages). *Gezin* 3:244-58.
- Weakliem, D. L. 1992. "Does Social Mobility Affect Political Behavior?" *European Sociological Review* 8:153-65.



# **Delinquency and Disdain: Social Capital and the Control of Right-Wing Extremism among East and West Berlin Youth<sup>1</sup>**

John Hagan  
*University of Toronto*

Hans Merkens  
*Freie Universität Berlin*

Klaus Boehnke  
*Technische Universität Chemnitz-Zwickau*

The authors link the notion of subterranean traditions to the concepts of control theory, anomic aspirations, and social capital to explain right-wing extremism and school delinquency among German youth. Weakened informal social controls and anomic aspirations lead to delinquent drift and extremist and delinquent involvements. East Berlin youth are uniquely exposed and vulnerable to anomic aspirations and associated right-wing extremism, but their schools and parents play significant roles in suppressing their right-wing attitudes. Schools and families are underappreciated sources of informal social control and resulting social capital that constrain right-wing extremism and related problems of young people during a period of rapid social change in the former East Germany.

Deviant subcultures are a source of fascination as well as fear, because while they characteristically stand in opposition to conventional culture, they nonetheless are connected to it. Obscured aspects of conventional society often are revealed by exploring the connections of subcultures to historical events and themes in the dominant culture through what Matza (1964, p. 63) calls a "subterranean analysis" of the "ongoing dialectic among a variety of conventional and deviant viewpoints." A key part

<sup>1</sup> This research was supported by a Killam Research Fellowship and grants from the German Research Council (DFG; 733/6-1,2 and 929/3-1,2,3). The assistance of Gabriele Classen in preparing the data for analysis is gratefully acknowledged. Correspondence may be directed to John Hagan, Faculty of Law, University of Toronto, 78 Queen's Park, Toronto M5S 2C5, Canada.

of this analysis involves the identification of subterranean traditions—in other words, practices and beliefs that are part of the mainstream culture but generally unacknowledged as such because they are also publicly condemned—and the processes that lead individuals to avoid or affiliate with them. This article focuses on the role of informal social control and the formation of social capital in restraining East and West Berlin youth from involvement in subterranean traditions of delinquency and right-wing extremism.

#### SUBCULTURES OF DELINQUENCY AND DISDAIN

Matza (1964 p. 64) identifies the irony of acceptance that distinguishes a subterranean tradition from other subcultural practices: "It is deviant, which is to say that it is publicly denounced by authorized spokesmen. However, the tradition is viewed with ambivalence in . . . privacy . . . and . . . the spirit and substance of subterranean traditions are familiar and within limits tolerated by broad segments of the adult population." By emphasizing the overlap of subcultural and cultural themes, Matza's concept of subterranean traditions avoids some of the empirical difficulties (see Kornhauser 1978) of more polarized statements about subcultures. Broadened versions of the perspectives on subcultures that incorporate overlap of subcultures with the dominant culture play an important role in contemporary understandings of racial or ethnic (Curtis 1975), class (Coleman 1989), and regional (Reed 1972, 1982) variations in attitudes and behaviors. Meanwhile, the convergence of subcultural and cultural themes may also be extended by an emphasis on entrepreneurial ambition and innovation in advanced capitalist societies.

Matza and Sykes (1961) illustrate this point by citing Thorsten Veblen's observation that delinquents conform to the norms of conventional society's business sector, rather than deviate from it, when they place a desire for "big money" in their value system. As Merton and others (e.g., Merton 1938; Bell 1953; Cloward and Ohlin 1960; Messner and Rosenfield 1994) also note, Western capitalism often elevates aspirations to wealth above norms of honesty and propriety. Matza and Sykes suggest that little more than a loosening of informal social control is required to unleash these aspirations, especially among adolescents. Thus while Merton and others postulate that great strains induce deviant adaptations to the wealth wish, Matza and Sykes believe that a free-wheeling and ambivalent culture presents a facilitative environment in which less constrained adolescents often simply drift toward marginally deviant innovations and actions (cf. Moscovici 1976).

Subcultural deviance does not merely take mild or material forms, however. Matza and Sykes (1961) also note that violence and aggression

are more accepted in our society than commonly is acknowledged. They observe that the "dominant society exhibits a widespread taste for violence, . . . in books, magazines, movies, and television. . . . The delinquent simply translates into behavior those values that the majority are usually too timid to express" (p. 716). This aggression is often mixed with disdain, as in the "actual use of aggression and violence in war, race riots, industrial conflicts, and the treatment of delinquents themselves by the police" (p. 717). Disdainful aggression can therefore also receive subterranean cultural support.

Matza and Sykes go on to suggest that the activist and wealth-motivated (some would say entrepreneurial) traditions of advanced capitalist culture encourage adventure, excitement, and thrill-seeking, which seemingly further promote deviance when compared with such conformity-producing values as security, routinization, and stability. The point is that the former, latent values exist side by side with the latter, more conventional values. Matza and Sykes (1961, p. 717) conclude that the "delinquent has picked up and emphasized one part of the dominant value system, namely, the subterranean values that coexist with other, publicly proclaimed values possessing a more respectable air."

These ideas have been explored mostly in North American society (e.g., Hagan 1991), but their explanatory power should be generic and therefore relevant to our understanding of deviance in other national settings as well. Indeed, as North American culture becomes more globally entwined—for example, with developments in Western and, more recently, Eastern Europe—we should observe a mix of historical and cultural themes that is likely to be varied in content but generic in process (DiMaggio and Mohr 1985). Familiar but understudied examples include transfers of punk-rock styles and skinhead movements from Western Europe to North America (see also Willis 1990). This article explores related subcultural delinquency and political processes using longitudinal survey data gathered in East and West Berlin.

#### A SOCIOLOGY OF DEVIANT POLITICAL BELIEFS

The collapse of communism and the embrace of market economics symbolized by the destruction of the Berlin Wall produced an unprecedented context for youths experiencing the transition to adulthood in East and West Berlin, as well as in other parts of eastern and central Europe. It is difficult if not impossible to know how much events surrounding the fall of the Berlin Wall intensified historical processes leading to the adoption of deviant political beliefs among German youth. Demonstrations of Nazi symbols and salutes preceded the recent transformative events in Germany, and historical forces involved in right-wing politics and associ-

ated hostilities toward foreigners were perhaps poised to resurface apart from German unification and the rejection of communism (Stüß 1993). These were latent political traditions awaiting a new generation of sub-cultural adherents. The collapse of communism and the embrace of market economics, however, likely added shape and direction to a more endemic cultural process.

In a revealing mix of the judicial and political, the former German Democratic Republic (GDR) sought to suppress the implications of right-wing extremism by defining the related acts of its young people as "hooliganism." Nonetheless, right-wing attitudes were expressed by the youth of East Germany throughout the 1980s, perhaps partly in rebellion against the communist regime, and since unification the former East Germany has experienced relative increases in conventional youth crime (albeit to levels that may still be lower than in the West) as well as highly publicized adolescent expressions of right-wing extremism and involvement in hate crimes, including violence against foreigners (Sessar 1991; Savelsberg 1995). Thus a government report estimates that as many as 30,000 Germans in the new states are involved in militant nationalist politics ("East Germany" 1990, p. 623), and survey research reveals that there are signs of substantial (though often unstable and organizationally and ideologically incoherent) right-wing extremism (Stöss 1991).

However, the combination of recent political events and historical and cultural forces in the new Germany suggest that the dismay expressed in some media reports about these events may be sociologically naïve (see also Childs 1985). A more useful analysis can link Matza and Sykes's (1961) conceptualization of subterranean traditions with Everett Hughes's (1964) classic analysis of "Good People and Dirty Work" in Nazi Germany. Hughes's essay parallels a modern control theory of crime (Hirschi 1969) that is linked to the work of Matza and Sykes. This theoretical perspective is anticipated when, rather than ask of Nazi Germany, "How did racial hatred rise to such high levels?" Hughes instead asks, "How could such dirty work be done among and, in a sense, *by* the millions of ordinary, civilized German people?" (1964, pp. 24–25; Hughes's emphasis). This reformulation parries the question commonly asked in psychologically oriented theories of crime and deviance, "Why do they do it?" with the question more provocatively posed by modern control theory: "How is it that more of us do not become involved?" (see Hirschi 1969). Hughes's anticipation of modern control theory suggests that citizens at large are more open to subterranean beliefs and behaviors than commonly is acknowledged in accounts of conventional culture.

Of course, Hughes's question involves not only those who participate in the behaviors under consideration but also those who allow these behaviors to happen. In addressing the latter group, Hughes focuses,

like Matza and Sykes, on currents of thought that often lie just below the surface of conventional society. Hughes (1964, p. 26) notes that the "essential underlying sentiments on racial matters in Germany were not different in kind from those prevailing throughout the western, and especially the Anglo-Saxon, countries." Hughes observes that such sentiments are common and that it takes relatively little to bring these sentiments to the surface of a national culture. The more interesting questions therefore ask how and by whom these sentiments can be constrained.

As in modern control theory, Hughes is concerned that across societies there are so many ready candidates and that "it would not take a very great turn of events to increase the number of such people, and to bring their discontents to the surface" (1964, p. 34). Hughes advocates a vigilant attitude by institutions of social control, including schools and the family, toward nascent groups with "negativistic and punishing attitudes." "It is a matter," writes Hughes (p. 35), "of checks and balances, of what we might call the social and moral constitution of society." In more contemporary terms, the Hughesian argument is that society must use mechanisms of social control and resulting social capital to constrain and contain the expression of subterranean cultural themes.

Hughes would no doubt be concerned about contemporary events in Germany. His perspective, combined with Matza and Sykes's focus on subterranean traditions and drift, gains relevance from the new circumstances. The economic challenges and the changing social and political order of a unified Germany raise concerns about the reemergence of subterranean traditions of right-wing extremism. How will individuals drift into these traditions? Can we identify generic social processes that constrain the direction taken by these events? How might the varied social contexts of East and West Germany condition these processes? Among the interesting answers to these questions is the possibility that East Germans even more than West Germans are drawing on underappreciated family- and school-based sources of social control and the resulting social capital to resist the drift of their youth into delinquency and right-wing extremism.

#### A MODEL OF YOUTHFUL RIGHT-WING EXTREMISM

Glen Elder (1974) observes that incongruencies of person and environment during crisis situations offer unique opportunities to study relationships between social structure and personality. A lesson of such research is that crises are not separately located in individuals, groups, or situations, but rather in their interaction, that is, in the interface of individuals, groups, and situations. Elder writes that "a crisis situation thus

refers to a type of *asynchrony* in the relationship between person or group and the environment" (p. 10; Elder's emphasis).

In this section we outline a theoretical perspective on the asynchronous situation confronting youth and families in the reunified East and West Berlin. It is important to emphasize first that the citizens of this previously divided city obviously have much in common. Citizens of the two Germanys were and are connected through literally millions of family ties and acquaintances and, perhaps as significantly, by predominantly watching the same West German television. Even before unification, then, most East Berlin youth were exposed along with West Berlin youth to the central themes of advanced capitalist culture and its subterranean traditions.

Our model therefore treats the postunification environments of East and West Berlin as both joined and residually distinct settings in which we can explore family and school control of involvement in the subterranean traditions of juvenile delinquency and right-wing extremism. Our initial expectation is that in both settings adolescent hostility toward foreigners and right-wing attitudes form internally coherent subterranean traditions that parallel traditions of juvenile delinquency. Like juvenile delinquency, these adolescent sentiments may be intensified by anomic aspirations. That is, anomic aspirations have been found in the past to be associated with juvenile delinquency (e.g., Farnworth and Leiber 1989), and we expect that this association will also be found with right-wing extremism, especially in East Berlin, as discussed below.

Unconstrained aspirations, anomic and otherwise, are expected to lead to these traditions through delinquent dispositions, or through a process of drift toward delinquency that involves spending time "hanging out" and "cruising" places that expose and attract the youth of both the East and the West to subterranean social and political traditions. These places are notably separate from the dominant controlling environments of the school and family and involve nonspecific teenage activities that lack explicit purpose or direction; a distinguishing characteristic of these activities and localities is that they place youth at risk for subterranean and nonconformist behaviors and beliefs.

Since schools are in most societies the central institution that monitors the transition from childhood to adulthood, they are a (perhaps the) classic target for expressions of adolescent autonomy and rebellion (Cohen 1955; Stinchcombe 1964). School vandalism and violence are therefore generic and commonly studied forms of juvenile delinquency across societies and cultures. Our point is that the same generic processes that lead to such delinquency also lead in the new German context (although not in this context alone) to traditions of right-wing politics and hostility toward foreigners.

Meanwhile, we also expect that adolescent inclinations to drift into subterranean traditions of right-wing extremism as well as juvenile delinquency will be constrained by informal social control processes involving success in school and family constraints that can assist the formation of social capital (Sampson and Laub 1993, p. 18; Coleman 1990, p. 305). The former processes are seen as generic social mechanisms that nominally seek to contain youth within publicly proclaimed normative spheres of social control, seeking, for example, to constrain otherwise anomic aspirations that may arise from the subterranean elevation of wealth and success above concerns about honesty and propriety. In this sense, school and family constraints are sources of social capital, resources that can be used to realize interests in avoiding undesired responses to asynchronous social situations.

In responding to the asynchrony of crisis situations, the above constraining processes represent "ways of dealing with resources and options that are employed in order to achieve control over the environment or life situation, to solve problems that arise from the disparity between claims and controls of outcomes" (Elder 1974, p. 10). This asynchrony likely increased under pressures of reunification in both West and East Berlin, even though for reasons we explore next, the interactions of individual circumstances and the surrounding situation were probably more consequential in the East.

To begin, youth in the East had less experience prior to unification with foreigners. Before unification, foreigners were a less visible part of East than West Berlin. Contract workers from other eastern block countries were not numerous, and they were a less prominent or visible part of East Berlin life. Following unification, foreign asylum seekers were relocated by law into East Berlin housing, often in recently vacated government buildings located in prominent parts of the city. Thus, while the presence of foreigners in West Berlin grew over a period of time, this presence in East Berlin was more sudden and associated with the other rapid changes occurring along with unification.

East Berlin families also confronted the greater challenge of reunification as they saw their city become a focal point for the integration of the new German economy and the monetary aspirations it raised but could not (at least in the short term) fulfill. By the early months of 1993, the previously uncommon experience of unemployment had become a normal part of the former GDR. Job security was much reduced, and rents and prices increased sharply (Savelsberg 1993).

Prior to this, the centrally planned and autocratic GDR also mandated a social structure that inevitably produced unintended as well as intended social differences that have enduring consequences in East German society. Borneman (1993) writes that, in a "process of mirror-imaging, the

two states fabricated themselves as moieties in a dual organization" (p. 17) and that "durable forms of division have . . . been built into the East-West distinction" (p. 334). Most conspicuously, of course, public life was strictly controlled in the former GDR. "By suppressing all other tendencies and developments," Rudolph (1987, p. 144) writes, "the GDR regime . . . succeeded in changing society" (p. 144). The rapid collapse of the GDR and the subsequent introduction of market economics and imagery left East Berlin youth and their families in classically anomic circumstances—caught between the repressive austerity of the old communism and the prospective riches of the new capitalism—and therefore vulnerably exposed to the anomie and related subterranean traditions we have described. Borneman (1993, p. 319) notes that "with the collapse of their state and social system, East Germans lost their time and space coordinates."

In the aftermath of these circumstances, there are some unique grounds for both pessimism and optimism regarding East German youth and their families. On the pessimistic side, we have already noted sound reasons to expect that adolescents in East Berlin would be more exposed and more vulnerable to anomic aspirations than adolescents in West Berlin, and they therefore could be expected to be more involved in subterranean traditions of right-wing extremism. The tendency of the former GDR to use antifascism as an ideology may also have generated right-wing extremism as a related subcultural response. However, on the more optimistic side, there are also reasons to believe that East German families and schools might provide significant informal social control, serving as sources of resulting social capital and acting to constrain the transmission of traditions of delinquency and right-wing extremism among their youth (Ewald 1991).

Coleman (1990) defines social capital as involving the creation of capabilities for action to achieve shared goals through socially structured relations between individuals in groups. When members of groups, for example in families and schools, are joined in the pursuit of common goals, Coleman describes them as creating social capital through the closure of social networks. Similarly, Sampson and Laub (1993) see social capital as deriving from strong social relations or social bonds in family, school, and other settings, and they note that, when these relations involve interdependence, they can provide resources that individuals can draw on in traversing difficult transitions in the life course (see also Braithwaite 1989; Hagan 1994, chap. 3).

Informal social control processes are therefore sources of social capital that can shield and protect youth from drifting into subterranean traditions of deviance and disorder during important transitional phases in the life course. Coleman (1990, p. 306) emphasizes that the current value



of the concept of social capital "lies primarily in its usefulness for qualitative analyses of social systems" rather than in its direct quantitative measurement. In this paper we directly measure differences in school performance and parental control of children in East and West Berlin, and we then draw inferences about resulting differences in the micro- and macrolevel formation of social capital in East and West Berlin. As Coleman (p. 305) notes, "the concept of social capital aids in both accounting for different outcomes at the level of individual actors and making the micro-to-macro transition without elaborating the social-structural details through which this occurs."

There are several reasons to expect that families of East Berlin youth could be especially significant sources of social control and resulting social capital relevant to subterranean traditions of juvenile delinquency and youthful right-wing extremism. For example, Elder (1974, p. 276) found that the family was an adaptive source of child control in his study of the Great Depression in Oakland, California. A more specific role of the family in reducing the occurrence of male delinquency during the Great Depression was revealed in studies in the United States (Glaser and Rice 1959) and Canada (McCarthy and Hagan 1987). The family can be a significant source of informal social control and the accumulation of social capital in economic, political, and social crises, and Coleman (1990, p. 593) writes that the "existence of a strong relation between adult and child must be regarded, on the whole, as a source of social capital important for the development of the child."

Historical and political forces noted above increased the salience of the family in East Germany. Although official policies supported collective, public forms of day care and early education that took young children out of the family at early ages, it was widely recognized that the family remained a uniquely influential institution in East German life. In a study that administered the Schwartz Value Survey (Schwartz and Bilsky 1990) to East and West German teachers and students studying to be teachers, Boehnke (1993) found that the value of "family security" ranked third among East Germans, while not ranking among the top 10 values among West Germans (see also Schlapentokh 1989, pp. 166–68).

Günter Gaus (1983, 1990), the former first representative of the Federal Republic of Germany in the GDR, contextualized the value attached to the privacy of family settings with the term *Nischengesellschaft*, or niche society. Similarly, Lemke (1989, p. 67) wrote just prior to unification that "families constitute the most important informal network for children and grown-ups. . . . the family represents a private niche, a realm of individual freedom and social relatedness that is characteristic for GDR society."

Recent accounts are explicit in emphasizing that this appreciation of

the privacy of the family in the GDR derived from the fact that the family was a niche of relatively unrestricted thought and expression within the repressive context of the GDR (Savelsberg 1993). Borneman (1993, p.321) concludes that "in the old GDR, public space had been so colonized by the state, so constrained in form, that any attempt to speak the self in public became an exercise in self-effacement. Thus people's primary investments were in private lives, private niches, private selves."

The residual strength and resilience of the East German family may continue to be an important institutional locus of social capital in the era of unification. For example, these family structures may be a source of informal social control of youthful drift toward the subterranean traditions discussed above. More specifically, the informal social control that flows from the East German families may constitute a significant resource that can help to restrain youthful inclinations to drift toward subterranean involvements in delinquency and right-wing extremism. Meanwhile, with the onset of adolescence, schools can become even more important than families as sources of informal social control and resulting sites for the formation of social capital.

The impact of schools on East German youth is not extensively studied. Early postwar efforts in East Germany officially focused on breaking the educational advantages of the old propertied classes (Childs 1988, p. 169). The educational testing necessary to fully assess the success of these efforts was rarely undertaken with GDR students. However, Borneman's (1993) recent interviews with postwar East Berliners suggest that the East German educational system was at least somewhat successful in diminishing economic class privileges. He notes that "in their life reconstructions, members of . . . working-class families tended to acknowledge that they benefited in their life trajectories from these affirmative action—or minimally, class-neutral—programs. Without the official programs, they said, they would never have had access to higher education" (p. 166).

Borneman (1993) particularly emphasizes the role of schooling in a mirror-imaging and differentiation of East and West Berlin that took place during the formative period of cold war comparisons. Contrasting themes of educational reform frequently were noted in the postwar period and were concentrated in West Germany around issues of talent, ability, and hierarchy. Borneman observes that during this period in West Germany, "the crucial concern of the schools was to reestablish hierarchy; that concern took the form of a debate over the nurturing of talent—a substance ascribed from birth and not learned. Schools with an open learning atmosphere, where different social classes mix, it was feared, might actually taint and spoil the inherited talent" (p. 242). This emphasis became an enduring feature of postwar education in the West. "The

official line was that 'resemblance' with the East in any respect was 'threatening' to the establishment of a distinct West German identity. In their reconstructions, . . . [respondents] frequently commented on three aspects of their education: its hierarchical nature, the pressure to be anti-communist, and the authoritarian nature of their teachers" (p. 242). It is perhaps unsurprising, then, that Borneman reports that many usually compliant youth in the West began to rebel against the schools and other symbols of state authority.

The schools in East Germany differed from those in the West in a number of ways: for example, by basing grades on group as well as individual accomplishments, by evaluating teachers as well as students on the basis of the grades the latter received, by involving students in assisting and tutoring their peers, by treating grades as a source of social support for students, and by more generally inflating the grades awarded. Grades were not regarded as measuring accomplishment or performance in the same sense as in the West. The emphasis was more collective and less individualistic. Students were less often segregated and stigmatized by grades, and they were integrated more into a collectively controlled classroom experience. Although some allowances must be made for the inevitable gap between ideological assertions and accomplishments, there was a growing respect for East German educational achievements among Eastern European states during the 1980s (Page 1985, p. 63). An ideological commitment to internationalism may also have helped to restrain hostilities toward foreigners and right-wing extremism among students.

So, while there may be much in need of change in the East German schools, it may also be that, by lifting and rewarding educational performance at lower levels, the East German schools were able to avoid some of the classic forms of delinquency directed against schools by frustrated students in the West, and they may also have played a role in constraining right-wing extremism that we have noted may otherwise be more characteristic of East Germany. Childs (1969, p. 199) ventured this kind of hypothesis when he ironically observed that, in spite of the tendency of GDR schools to accept and impose party ideology, "western educationalists . . . will admire the progress made in other aspects of education, progress which may well, in the long run, help to remove some of the less desirable features of East German society today." Within the theoretical framework we have proposed, East Berlin schools may be a significant source of informal social control, and resulting social capital, that restrains youthful inclinations to drift toward subterranean involvements in delinquency and right-wing extremism. In producing this control, schools may occupy an even more important role than families. This

does not discount the salience of the family at earlier ages; however, by adolescence the burden of influence may have shifted.

#### DATA AND MEASURES

We turn next to a description of the sample, measures, and methods we use to explore and test the theoretical framework we have proposed. The data analyzed in this paper come from a continuing panel study conducted with students who were in the seventh and ninth grades in 1991, and the eighth and tenth grades in 1992, in 16 public schools in West and East Berlin. The schools were selected following a detailed pilot study of parts of the city judged to be typical of East and West Berlin. Housing is more crowded and unemployment is higher in East Berlin than in West Berlin, and the communities from which schools were selected for this study closely reflect overall differences on these dimensions (see Kirchhöfer, Merkens, and Steiner 1991).

The panel consists of 489 students who continued in this study from 1991 through 1992. While exact enrollment figures are not available for several of the schools, the response rates across schools and sections of the city are estimated to be between 60% and 70%. The retention rate across waves was 72.4%. Proportions of students do not differ significantly between waves of the survey, by gender, across respective grades, and by types of schools, and these proportions are closely representative of the communities from which they are drawn (see Merkens, Kirchhöfer, and Steiner 1992).

The subterranean nature of the traditions that concern us suggests that they will be reflected among public school students. An estimated 70% of the individuals involved in officially recorded assaults against "guest workers" and refugees in 1992 were under 20 years of age, while only 10% were identified as skinheads, and fewer than 15% had any established link to neo-Nazi groups. These incidents characteristically involve individuals who are otherwise within the mainstream of German adolescents (Viviano 1993).

The longitudinal measurement of variables in the panel is used to reflect the sequence of the model we have presented. Aside from the exogenous variables (sex, age, and location in East or West Berlin), parental control, school achievement, and anomic aspirations were measured in 1991, while delinquent drift, involvement in school vandalism and violence, hostility toward foreigners, and right-wing orientations were measured in 1992. Alpha reliability scores, descriptive statistics, and difference of means tests are presented in table 1.

Hostility toward foreigners was measured by Likert-style responses to

TABLE 1

## DESCRIPTIVE STATISTICS FOR DEPENDENT AND INDEPENDENT VARIABLES

VARIABLES	EAST BERLIN			WEST BERLIN			t-Value of Difference
	$\alpha$	Mean	SD	$\alpha$	Mean	SD	
Hostility toward foreigners <sup>a</sup> .....	.80	.99	.58	.80	.85	.59	2.64*
Right-wing orientation <sup>a</sup> .....	.77	.73	.71	.65	.64	.63	1.48
School vandalism and violence <sup>b</sup> .....	.69	.48	.58	.76	.58	.66	-1.75
Delinquent drift <sup>a</sup> .....	.40	1.26	.55	.59	1.29	.53	-.62
Anomic aspirations <sup>a</sup> .....	.52	1.51	.63	.48	1.31	.61	3.57*
Parental control <sup>c</sup> .....	.69	.81	.28	.67	.66	.35	5.36*
School achievement <sup>a</sup> .....	.74	3.77	.73	.66	3.12	.74	9.70*
Sex <sup>d</sup> .....		1.57	.50		1.49	.50	1.78
Age .....		14.52	1.13		14.62	1.22	-.94
N .....		246			243		

NOTE.—Questions used to define composite scales are discussed in the "Data and Measures" section of the text. Results reported in the table have been coded such that higher scores indicate more of the quality being measured, i.e., more hostility, more right-wing orientation, more involvement in vandalism and violence, higher school achievement, etc. Mean scale scores were calculated for individuals and used in the analysis.

<sup>a</sup> On a scale from 0 to 4.

<sup>b</sup> Incidents coded as never = 0, once = 1, twice = 2, three times = 3, four or more times = 4.

<sup>c</sup> Items coded "0" or "1."

<sup>d</sup> Male = 1, female = 2.

\*  $P < .01$ , two-tailed test

six items asserting that "foreigners" are better off in their homelands, not as trustworthy as Germans, should be dismissed from jobs before Germans, and are too numerous in Germany, that foreign aid is overly philanthropic, and that Germany's first priority should be to restore its "earlier greatness." As expected, in table 1 we see that East Berlin youth score significantly higher ( $P < .01$ , two-tailed test) on this set of items than do West Berlin youth.

Right-wing orientations similarly were measured with Likert-style responses to the slogans, "Führer command, we will follow," "Germany—the only true future," "Smash the leftists," and "Keep Germany clean." The difference between East and West youth on this scale was in the same direction but weakly so (i.e., significant at the  $P < .10$  level, one-tailed test).

Involvement in delinquency was measured with responses from youth indicating how often they engaged in school vandalism and violence that included destroying school property, victimizing another person at school, and "thrashing" (i.e., beating up) others at school (responses ranged from never to four or more times). West Berlin youth scored marginally higher than East Berlin youth on these items ( $P < .10$ , two-tailed test).

Delinquent drift was measured from reports of the amount (from a little to all) of leisure time spent "hanging out," "on things other than school activities," "cruising," or "acting tough." East and West Berlin youth spent approximately the same amount of time in these ways.

Anomic aspirations were measured with Likert-style responses to statements asserting: "It is not important how you get money, you have to have money to get ahead"; "The most important thing in getting ahead is knowing and getting along with important people"; and "Winning is the most important thing in life, not how you win." Again as expected, East Berlin youth scored significantly higher ( $P < .01$ , two-tailed test) than West Berlin youth on these measures of anomie.

Our measures of informal social control linked to the formation of social capital in the family involve items reflective of parental control. These items asked if the youth surveyed could meet friends after 8:00 p.m., stay over at a male or female friend's home, or take a holiday trip without checking with parents. These measures are reflective of constraint within families. As expected, we found significantly ( $P < .01$ , two-tailed test) more informal social control within the families of East Berlin youth than West Berlin youth. This result is consistent with the emphasis on the family as a valued "private niche" in East German society.

Our measures of success in school (self-reported math and German grades and overall academic performance) are also linked to the formation of social capital. Grades are a measure of the acquisition of human capital, for example, in the form of technical skills, but grades also play a role in the formation of social capital, through the creation and certification of cultural capacities and credentials (see Robinson and Garnier 1985, pp. 253–55; also Menard and Morse 1984). Self-reported grades and overall ratings of performance (from poor to excellent) reflect the impression the surveyed youth have of their success in school. Sampson and Laub (1990, p. 611) note that the sense of "*social investment* or social capital in the institutional relationship" (emphasis in original) can determine the salience of controlling effects of institutional experiences such as schooling. East Berlin youth self-reported significantly ( $P < .01$ , two-tailed test) higher levels of school success than West Berlin youth; for reasons noted above, this group difference may result in the formation of a more collectively oriented sense of vested social capital, which may enhance the controlling effects of East Berlin schooling experiences.

Table 1 also provides information about the sex and age of the students surveyed. Although there are marginally more girls in the East Berlin sample than in the West Berlin sample, both groups are between 14 and 15 years of age, and about half of the sample lives in each part of Berlin.

TABLE 2  
LISREL MEASUREMENT MODELS FOR SCHOOL DELINQUENCY AND  
RIGHT-WING EXTREMISM

Model	$\chi^2$	df	GFI
A: Null model—items uncorrelated .....	2,818.07	78	.487
B: One-factor model .....	933.55	65	.826
C: Three-factor model—factors correlated .....	353.40	62	.931
D: One-factor model—correlated errors within subsets of items ... ..	111.63	41	.978
E: Two-factor model—factors correlated and correlated errors within subsets of items . ....	116.75	43	.976

NOTE —For all models  $P < .001$

### THE ANALYSIS

Alpha reliability scores are reported for the multi-item scales presented in table 1. These scores range from .40 to .80, predictably increasing with the number of items included (see Nunnally 1978, p. 223). Single factor LISREL measurement models also were estimated separately for each of the explanatory concepts and their specified measures. In each case this resulted in an acceptable model with a goodness-of-fit index (GFI) above .95 and a ratio of chi-square to degrees of freedom of less than two.

LISREL measurement models were also used initially to analyze jointly the indicators of school delinquency and right-wing extremism. The results of estimating several measurement models are presented in table 2. An initial null model A leaves the items that measure hostility toward foreigners, right-wing orientation, and school vandalism and violence uncorrelated, thereby providing a baseline comparison measure of model fit ( $\chi^2 = 2,818$ ;  $df = 78$ ;  $GFI = .487$ ). Simply treating these items as a single factor in model B significantly improves the fit with the data ( $\chi_A^2 - \chi_B^2 = 1,884.52$ ;  $df_A - df_B = 13$ ;  $P < .01$ ), raising the GFI from .487 to .826. A three-factor model with correlations among the factors significantly improves this fit ( $\chi_B^2 - \chi_C^2 = 580.15$ ;  $df_B - df_C = 3$ ;  $P < .01$ ), increasing the GFI to .931. However, this fit can be improved in a number of additional ways.

The best-fitting models either treat all of the items as representing one factor or two, with correlated errors allowed within subsets of items. These models, models D and E, are a significant improvement over model C (e.g.,  $\chi_C^2 - \chi_D^2 = 241.67$ ;  $df_C - df_D = 21$ ;  $P < .01$ ), both raising the GFI to over .97. The difference in the fit between Models D and E ( $\chi_E^2 - \chi_D^2 = 5.12$ ;  $df_E - df_D = 2$ ;  $P > .05$ ) is not significant,

TABLE 3  
STANDARDIZED MEASUREMENT COEFFICIENTS FOR MODIFIED TWO-FACTOR MODEL  
(MODEL E)

Item	Coefficient
School vandalism and violence:	
Destroyed school property . . . . .	.67
Victimized others at school . . . . .	.80
Trashed school property . . . . .	.69
Right-wing extremism:	
Foreigners to homelands . . . . .	.49
Foreigners untrustworthy . . . . .	.69
Dismiss foreigners . . . . .	.62
Foreigners bad for Germany . . . . .	.74
Foreign aid wasteful . . . . .	.58
Restore Germany's greatness . . . . .	.56
Follow führer . . . . .	.28
Germany's true future . . . . .	.56
Smash leftists . . . . .	.47
Keep Germany clean . . . . .	.56

NOTE.—Coefficients are for the combined sample of East and West Berlin youth. The correlation between factors is .33. Items are discussed in the "Data and Measures" section of the article

making the choice between these models moot. The standardized measurement coefficients from the two-factor model with errors correlated within subsets of items are presented in table 3; the correlation between the school vandalism and violence factor and the right-wing extremism factor is .33. These results are consistent with our expectation that the subterranean tradition of right-wing extremism includes internally coherent items that parallel school vandalism and violence among East and West Berlin youth. Since our interest is in further exploring similarities and differences in processes that lead to these subterranean traditions, we consider these traditions separately below.

In the structural models we estimate next with LISREL, therefore, we treat right-wing extremism as a latent construct observed through items measuring hostility toward foreigners and right-wing orientations, while we model school vandalism and violence as a separate factor. The results of fitting several structural models for the combined East and West Berlin subsamples are summarized in table 4.

We begin with a null model, A, in which no structural effects are postulated as a baseline for comparison. Next we estimate a pure mediation model, B, in which gender, age, and location in East or West Berlin exert their influence through the school- and parent-based measures of social control, as well as through anomic aspirations, in determining the



TABLE 4

LISREL STRUCTURAL MODELS FOR COMBINED EAST AND WEST BERLIN SAMPLES

Model	$\chi^2$	<i>df</i>	GFI	<i>P</i>
A: Null model—no effects .....	513.10	43	.813	<.001
B: Pure mediation model .....	152.53	29	.940	<.001
C: Mediation model and direct effect of gender on school vandalism and violence .....	100.71	28	.961	<.001
D: Fully modified mediation model .....	40.97	25	.984	=.023

delinquent drift that leads in parallel but separate ways to right-wing extremism and school vandalism and violence. This model B is a substantial improvement over model A ( $\chi_A^2 - \chi_B^2 = 360.57$ ;  $df_A - df_B = 14$ ;  $P < .01$ ). Nonetheless, an obvious omission from this model is the direct effect, which much work in sociological criminology (e.g., see Harris 1977) confirms, of gender on measures of delinquency that include school vandalism. When this effect is added in model C, the fit of the model is further improved ( $\chi_B^2 - \chi_C^2 = 51.28$ ;  $df_B - df_C = 1$ ;  $P < .01$ ).

Examination of the modification indices associated with model C led us to make several further changes that make theoretical sense in estimating the final model, D, whose structural effects are summarized in figure 1. The most important changes in theoretical terms involve adding direct effects of living in East Berlin on anomic aspirations and of having anomic aspirations on right-wing extremism. Further direct effects are added from age and school achievement to right-wing extremism, while a path from age to school achievement is deleted. Although these changes were suggested by modification indices, they are all consistent with the theoretical model developed earlier. The resulting model D ( $\chi_C^2 - \chi_D^2 = 59.74$ ;  $df_C - df_D = 3$ ;  $P < .01$ ) is further improved, with a GFI of .984, and a highly satisfactory ratio of chi-square to degrees of freedom (1.64; see Carmines and McIver 1981). The modification index for the correlation of errors between school vandalism and violence and right-wing extremism is below the conventional level that encourages revision and no further changes are indicated in the model.

The standardized structural effects resulting from estimation of model D are summarized in figure 1 for the combined East and West Berlin subsamples. Separate East and West subsample coefficients also are presented respectively in parentheses, with an asterisk following when the difference in coefficients between subsamples is significant at the .05 level.

All the effects in the combined subsamples are significant at the .01

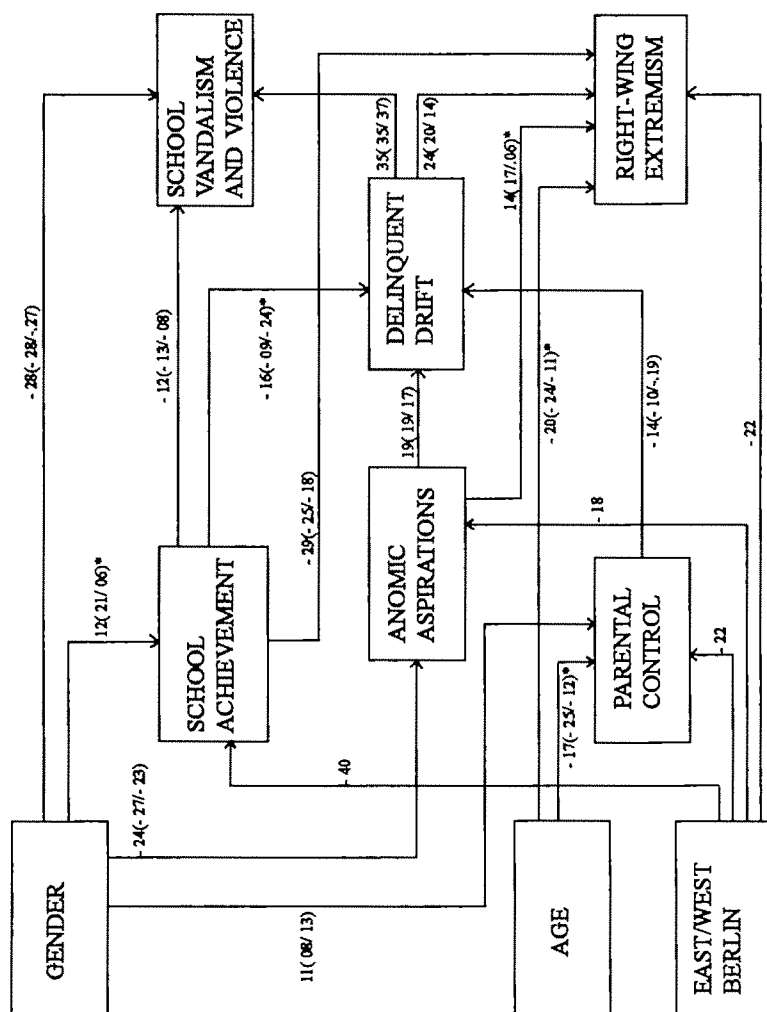


FIG. 1.—Path model for school delinquency and right-wing extremism for all Berlin youth (East and West Berlin youth separately in parentheses). Loadings for hostility toward foreigners and for right-wing orientation as measures of right-wing extremism are .81 and .73, respectively. Asterisk indicates difference between paths is significant at the .05 level, two-tailed test.

level and beyond, and these effects generally support the theoretical perspective introduced above. For example, recall our expectation that anomic aspirations would intensify school vandalism and violence as well as right-wing extremism. Anomic aspirations indirectly increase school vandalism and violence as well as right-wing extremism by increasing delinquent drift. That is, as expected, more anomic youth are more likely to spend time in leisure settings ( $\beta = .19$ ) that increase exposure to traditions of delinquency and right-wing extremism, thereby increasing involvement in delinquent ( $\beta = .35$ ) and extremist ( $\beta = .24$ ) traditions. The separate subsample estimates reveal that anomic aspirations also directly increase right-wing extremism in East Berlin ( $\beta = .17$ ). The latter finding is consistent with our expectation that youth in East Berlin are uniquely vulnerable to the effects of anomic aspirations.

There is also evidence in figure 1 that adolescent inclinations to drift into subterranean traditions of right-wing extremism as well as juvenile delinquency are constrained by informal social controls associated with the formation of social capital; this formation of social capital involves self-reported academic success in school and informal social control within the family. These constraining influences are reflected in the respective indirect effects of school achievement ( $\beta = -.16$ ) and parental control ( $\beta = -.14$ ) operating through delinquent drift, as well as through the direct effect of school achievement on school vandalism and violence ( $\beta = -.12$ ) and on right-wing extremism ( $\beta = -.29$ ). Interestingly, the separate subsample estimates reveal that the effect of school achievement on delinquent drift is unique among West Berlin youth ( $\beta = -.24$ ). This may reflect a reaction to the more hierarchical tendencies of the West German school system that we discussed above.

However, perhaps the most revealing results in figure 1 are with regard to the main effects of youth living in East and West Berlin. Our theoretical perspective reflects contemporary speculation about events in Germany in predicting that adolescents in East Berlin are more exposed and more vulnerable to anomic aspirations than adolescents in West Berlin and that they therefore identify more strongly with subterranean traditions of right-wing extremism. We have already seen that youthful anomic aspirations have the expected direct effect on right-wing extremism in East Berlin. The structural paths in figure 1 also indicate that living in East Berlin increases identification with right-wing extremist views directly ( $\gamma = -.22$ ) as well as indirectly by increasing anomic aspirations ( $\gamma = -.18$ ). The former direct effect is actually stronger than the difference of means reported in table 1 suggests. The implication is that this effect is suppressed by countervailing forces.

These countervailing forces involve the East German families and schools discussed earlier. Our perspective suggested these institutions

might provide important sources of informal social control and facilitate the formation of social capital that can limit the transmission of traditions of delinquency and right-wing extremism by helping to restrain inclinations of East German youth to drift toward these subterranean involvements. The indirect means by which this occurs involve the higher levels of parental control ( $\gamma = -.22$ ) and school achievement ( $\gamma = -.40$ ) reported by East Berlin youth. While the indirect effect of parental control in reducing school delinquency and right-wing extremism in East German families is notable, the role of school achievement is especially significant.

More specifically, the indirect role of school achievement in reducing right-wing extremism plays a predominant part in suppressing the effect of living in East Berlin on right-wing extremism. As indicated above, the direct effect of living in the East is  $-.22$ . However, the total effect of living in the East on right-wing extremism is suppressed in a countervailing fashion by the indirect effect of the higher reported levels of school achievement among those living in the East ( $-.40 \times -.29 = .12$ ). This indirect effect reduces the total effect of living in the East by nearly half. The effect of parental control operating through delinquent drift reduces this total effect slightly further (i.e., by  $-.22 \times -.14 \times .24 = .007$ ). The implication is that right-wing extremism among East Berlin youth would be much more prevalent if the influence of schools and families in the East did not constrain identification with extremist beliefs.

Finally, it is of interest to note that the effects of age on parental control and right-wing extremism are greater in East Berlin than in West Berlin. Both decline more quickly with age in the East than West. Although not anticipated in our theoretical framework, this result suggests that the East is a more age-graded setting. Gender is also more strongly related to school achievement in East Berlin, suggesting that gender as well as age stratification is more intense in the East.

## DISCUSSION AND CONCLUSIONS

The analysis presented in this paper indicates that the subterranean traditions of right-wing extremism and school delinquency parallel one another among Berlin youth. This parallelism is reflected in the approximately equal fit of measurement models that treat indicators of right-wing extremism and school delinquency as combined or separate factors. It is further reflected in the structural models that portray similar developmental sequences that lead from weaknesses of informal social control and anomic aspirations through delinquent drift to extremist and delinquent involvements. Although there are differences across settings and outcomes as well, the traditions of adolescent right-wing extremism

and school-based juvenile delinquency exhibit striking parallels in East and West Berlin, raising the possibility, in need of further test, that this similarity is present in other societal settings as well. The extension of this model to the parallel study of the skinhead movement and crime and delinquency in North America is one possible example.

The role of anomic aspirations in the development of right-wing extremism in contemporary Germany is of special interest. Our findings indicate that anomic aspirations and right-wing extremism are, as expected, higher among the youth of East Berlin than of West Berlin and that East Berlin youth are especially vulnerable to the effects of anomic aspirations. This difference suggests that it is not only relaxation of social control that leads to the mobilization of an unrestrained "wealth wish" among adolescents, but that repression and deprivation have also played key roles in intensifying anomie in the East. This of course implies that it is the presence of strain as well as the absence of control that increases anomic aspirations in the development of right-wing extremism. These findings provide a new source of evidence about the unsettled influence of strain in deviance and social problems theory (see, e.g., Kornhauser 1978).

Meanwhile, the role of informal social control in constraining drift toward traditions of right-wing extremism suggests that insights of modern control theory can be extended beyond the explanation of crime and delinquency to include extreme political beliefs and movements as well. It may bear note that this role also includes a path of influence that leads from gender through parental control and delinquent drift to right-wing extremism and school vandalism and violence. This path of influence is consistent with the recent development of a power-control theory that postulates gender differences in parental control as a significant source of variation in social deviance (see Hagan 1989; Grasmick, Blackwell, and Bursik 1993).

However, our analysis also reveals that the intervening role of school achievement is more salient than parental controls in explaining differences in right-wing extremism in East and West Berlin. School achievement is especially salient in suppressing what would otherwise be a substantially larger difference in the right-wing extremism of East and West Berlin youth. Families may play a bigger role at earlier ages, but schools seem more important in the early years of German adolescence. This will make it particularly important to consider how ongoing reforms in the East German schools might affect their restraining effects. Both schools and families may be underappreciated sources of informal social control and the formation of social capital that are helpful in containing right-wing extremism and related problems of young people during a period of rapid change in East Germany.

It is important in the context of this paper to keep in mind the explicit premise of modern control theory that the constraining influences of institutions such as the family and school can be put to varied moral purposes (Hirschi 1969). It is likely that the family and schools in Nazi Germany were involved as institutions of social control in encouraging conformity to attitudes and actions that from the perspective of a larger global community and many inside Germany were linked to extreme forms of deviance. The point is not that these controls are good or bad, in themselves, but rather that they are influential in channeling social behavior in varying directions. The lesson of Nazi Germany is that in some historical circumstances subterranean traditions (e.g., anti-Semitism) that exist alongside other more prominently ambivalent societal themes (e.g., German nationalism) can become ascendant and become a part of conformity-producing institutions and processes of social control. Social control is understood in modern control theory as a potential source of social evil as well as social good. The variability of the purposes to which social control is put helps to account for what Hannah Arendt appropriately has called the "banality of evil."

In considering the role of schools and families in the future of East Germany, it may be significant to recall the use that Elder (1974, p. 295) makes of Sherif's (1958) notion of a superordinate goal. In the current case, a superordinate goal includes the challenge of maintaining a social fabric that is seriously threatened, for example, by right-wing extremism, during the transition from a command to a market economy. The findings of this research suggest that the schools and families of East Germany have residual strengths that are important resources for this task. It may be crucial to find ways of maintaining and coordinating the social capital that can be derived from these changing institutions along with the infusion of financial capital from West Germany during the unified pursuit of economic and national goals. This joining of social and economic resources may be a key part of putting aside internal differences in the pursuit of a larger national purpose. Alternatively, our findings suggest that ignoring this possibility would constitute a lost opportunity for successful social change.

Finally, this research illustrates the contribution that comparative longitudinal research can play in the development of what Sampson (1993) has called a dynamic contextual approach to the study of crime and social deviance in the life course. The results of this study do this by underlining the roles that location in time and place play in shaping the expression of social deviance. Sociologists of crime and deviance have been surprisingly slow to take the study of their subject matter out of the contemporary United States, and as a result we know less about the contextual and dynamic features of these phenomena than we should. For example,

this analysis raises questions to be answered in further research about the comparative and longitudinal effects of social capital and anomie, as well as forms of social deviance including school delinquency and right-wing extremism, in shaping longer-term life outcomes in East and West Berlin and beyond.

## REFERENCES

- Bell, Daniel. 1953. "Crimes as an American Way of Life." *Antioch Review* 13. 131-54.
- Boehnke, Klaus. 1993. "Lehrerinnen und Lehrer als Wertemultiplikatoren im veränderten Bildungssystem der neuen Bundesländern: Probleme und Perspektiven" (Teachers as value multipliers in the changing educational system of the new states: Problems and perspectives). *Pädagogik und Schulalltag* 48:94-105.
- Borneman, John. 1993. *Belonging in the Two Berlins: Kin, State, Nation*. Cambridge: Cambridge University Press.
- Braithwaite, John. 1989. *Crime, Shame, and Reintegration*. Cambridge: Cambridge University Press.
- Carmines, E. G., and J. P. McIver. 1981. "Analyzing Models with Unobserved Variables: Analysis of Covariance Structures." Pp. 65-116 in *Social Measurement: Current Issues*, edited by George Bohrnstedt and Edgar Borgatta. Beverly Hills, Calif.: Sage.
- Childs, David. 1969. *East Germany*. New York: Praeger.
- . 1985. "Youth—Not So Very Different." Pp. 118-33 in *Honecker's Germany*, edited by David Childs. London: Allen & Unwin.
- . 1988. *The GDR: Moscow's German Ally*. London: Unwin Hyman.
- Cloward, Richard, and Lloyd Ohlin. 1960. *Delinquency and Opportunity*. Glencoe, Ill.: Free Press.
- Cohen, Albert. 1955. *Delinquent Boys*. Glencoe, Ill.: Free Press.
- Coleman, James. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Coleman, James W. 1989. *The Criminal Elite*. New York: St. Martin's.
- Curtis, Lynn. 1975. *Violence, Race, and Culture*. Lexington, Mass.: Heath.
- DiMaggio, Paul, and John Mohr. 1985. "Cultural Capital, Educational Attainment, and Marital Selection." *American Journal of Sociology* 90:1231-61.
- "East Germany Sets October 3 Reunification Date." 1990. *Facts on File World News Digest* 50 (August 24): 622-23.
- Elder, Glen. 1974. *Children of the Great Depression: Social Change in the Life Experience*. Chicago: University of Chicago Press.
- Ewald, Uwe. 1991. "The Social Transformation Process in the Former German Democratic Republic with Respect to Crime." Paper presented at the American Society of Criminology meetings, San Francisco.
- Farnworth, M., and M. J. Leiber. 1989. "Strain Theory Revisited: Economic Goals, Educational Means, and Delinquency." *American Sociological Review* 54:263-74.
- Gaus, Günter. 1983. *Wo Deutschland liegt: Eine Ortsbestimmung* (Where Germany is situated: A definition of place). Munich: DTV Zeitgeschichte.
- . 1990. *Deutscher Lastenausgleich* (Equalization of burdens in Germany). Frankfurt: Luchterhand.
- Glaser, Daniel, and Kent Rice. 1959. "Crime, Age, and Employment." *American Sociological Review* 24:679-84.

- Grasmick, Harold, Brenda Blackwell, and Robert Bursik. 1993. "Changes in the Sex Patterning of Perceived Threats of Sanctions." *Law and Society Review* 27:679-706.
- Hagan, John. 1989. *Structural Criminology*. New Brunswick, N.J.: Rutgers University Press.
- . 1991. "Destiny and Drift: Subcultural Preferences, Status Attainments, and the Risks and Rewards of Youth." *American Sociological Review* 56:567-82.
- . 1994. *Crime and Disrepute*. Newbury Park, Calif.: Pine Forge Press.
- Harris, Anthony. 1977. "Sex and Theories of Deviance: Toward a Functional Theory of Deviant Type-Scripts." *American Sociological Review* 42:3-16.
- Hirschi, Travis. 1969. *Causes of Delinquency*. Berkeley: University of California Press.
- Hughes, Everett. 1964. "Good People and Dirty Work." Pp. 23-36 in *The Other Side. Perspectives on Deviance*, edited by H. Becker. New York: Free Press.
- Kirchhöfer, Dieter, Hans Merckens, and Irmgard Steiner. 1991. *Schuljugendliche im vereinten Berlin: Pilotstudie* (School youths in the united Berlin: A pilot study). Berlin: Freie Universität Berlin.
- Kornhauser, Ruth. 1978. *The Social Sources of Delinquency*. Chicago: University of Chicago Press.
- Lemke, Christiane. 1989. "Political Socialization and the 'Micromilieu': Toward a Political Sociology of GDR Society." Pp. 59-73 in *The Quality of Life in the German Democratic Republic*, edited by Marilyn Rueschemeyer and Christiane Lemke. London: M. E. Sharpe.
- McCarthy, Bill, and John Hagan. 1987. "Gender, Delinquency, and the Great Depression: A Test of Power-Control Theory." *Canadian Review of Sociology and Criminology* 24:153-77.
- Matza, David. 1964. *Delinquency and Drift*. New York: Wiley.
- Matza, David, and Gresham Sykes. 1961. "Juvenile Delinquency and Subterranean Values." *American Sociological Review* 26:712-20.
- Menard, Scott, and Barbara Morse. 1984. "A Structuralist Critique of the I.Q.-Delinquency Hypothesis: Theory and Evidence." *American Journal of Sociology* 89 (6): 1347-78.
- Merckens, Hans, Dieter Kirchhöfer, and Irmgard Steiner. 1992. *Berliner Schülerstudie 1992* (Berlin student study 1992). Berlin: Freie Universität Berlin.
- Merton, Robert. 1938. "Social Structure and Anomie." *American Sociological Review* 3:672-82.
- Messner, Steven, and Richard Rosenfeld. 1994. *Crime and the American Dream*. Belmont, Calif.: Wadsworth.
- Moscovici, Serge. 1976. *Social Influence and Social Change*. London: Academic Press.
- Nunnally, J. C. 1978. *Psychometric Theory*. New York: McGraw-Hill.
- Page, John. 1985. "Education under the Honeckers." Pp. 50-65 in *Honecker's Germany*, edited by David Childs. London: Allen & Unwin.
- Reed, John. 1972. *The Enduring South: Subcultural Persistence in Mass Society*. Lexington: Heath.
- . 1982. *One South: An Ethnic Approach to Regional Culture*. Baton Rouge: Louisiana State University Press.
- Robinson, Robert, and Maurice Garnier. 1985. "Class Reproduction among Men and Women in France. Reproduction Theory on Its Home Ground." *American Journal of Sociology* 91 (2): 250-80.
- Rudolph, Herman. 1987. "The Human Dimension." Pp. 141-56 in *Germany between East and West*, edited by Edwina Moreton. Cambridge: Cambridge University Press.
- Sampson, Robert. 1993. "Linking Time and Place: Dynamic Contextualism and the Future of Criminological Inquiry." *Journal of Research in Crime and Delinquency* 30:426-44.



- Sampson, Robert, and John Laub. 1990. "Stability and Change in Crime and Deviance over the Life Course: The Salience of Adult Social Bonds." *American Sociological Review* 55:609-27.
- . 1993. *Crime in the Making: Pathways and Turning Points through Life*. Cambridge, Mass.: Harvard University Press.
- Savelsberg, Joachim. 1993. "Crime and Political Constitution: The East Central European Transition to Democracy." Paper presented at the American Sociological Association meetings, Miami, August.
- . 1995. "Crime, Inequality, and Justice in Eastern Europe: Anomie, Dominance, and Revolutionary Change." In *Crime and Inequality*, edited by John Hagan and Ruth Peterson. Stanford, Calif.: Stanford University Press. In press.
- Schlapentokh, Vladimir. 1989. *Public and Private Life of the Soviet People: Changing Values in Post-Stalin Russia*. New York: Oxford University Press.
- Schwartz, Shalom H., and Wolfgang Bilsky. 1990. "Toward a Theory of the Universal Content and Structure of Values: Extensions and Cross-Cultural Replications." *Journal of Personality and Social Psychology* 58:878-91.
- Sessar, Klaus. 1991. "Crime Rate Trends before and after the End of the German Democratic Republic." Paper presented at the American Society of Criminology meetings, San Francisco, November.
- Sherif, Muzafer. 1958. "Superordinate Goals in the Reduction of Intergroup Tensions." *American Journal of Sociology* 53:349-56.
- Stinchcombe, Arthur. 1964. *Rebellion in a High School*. Chicago: Quadrangle Books.
- Stöss, Richard. 1991. "Right-Wing Extremism in East and West Germany." Paper presented at the Conference on the Radical Right in Western Europe, Minneapolis.
- Süß, Walter. 1993. "Was wußte die Stasi über die Neonazis in der DDR?" (What did the Stasi know about neonazis in the GDR?). *Die Zeit*, no. 18, p. 40.
- Viviano, Frank. 1993. "Germany's Junk-Strewn Heart of Darkness." *Globe & Mail* (July 10): A1.
- Willis, Paul. 1990. *Common Culture*. Boulder, Colo.: Westview.

## Book Reviews

*The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century.* By Kim Voss. Ithaca, N. Y.: Cornell University Press, 1993. Pp. xv + 290. \$39.95 (cloth); \$16.95 (paper).

Jeffrey Haydu

*University of California, San Diego*

One reason why the debate over "American exceptionalism" persists is the imprecision of the concept. What is so peculiar about U.S. class relations? Is it workers' limited solidarity? The weakness of leftist political movements? Labor's narrow political vision? The anticollectivist character of industrial relations? And what countries set the standard for "normal" class relations? Kim Voss's lucid and important book zeros in on working-class solidarity—specifically, institutionalized alliances between craftsmen and less skilled workers—and makes England and France the comparative reference points. Having defined "exceptionalism" in this way, Voss then traces its origins to the demise of the Knights of Labor in 1886–87.

Earlier in the 19th century, Voss reminds us, labor movements in England, France, and the United States had narrow artisanal roots. Only by the late 1800s did changes in work and in urban communities make possible much more broadly based working-class movements. Here Voss joins other students of class formation in rejecting traditional assumptions that working-class solidarity is in any country the "normal" outcome of capitalist development. Instead, she highlights divisions among workers and asks how, if at all, different sectors of the working class developed and consolidated alliances. In the Knights' heyday (1884–86), American workers, like their counterparts in England and France at about the same time, overcame their differences.

Both structural innovations (particularly the range of organizational forms permitted at the local level) and ideological appeals (especially the vision of equal citizenship in an industrial commonwealth purged of parasitical monopolists) brought workers of varied skills and interests into the Knights. To more fully explain the movement, Voss brings to bear an impressive database (painstakingly assembled from a wide range of archival and secondary sources) on New Jersey locals and the industrial and community settings in which they appeared. This collection of data is one of the major strengths and contributions of the book, allowing the author to evaluate the roots of solidarity with statistical precision. Humanists need not despair, however. Voss complements each major

Permission to reprint a book review printed in this section may be obtained only from the author.

argument with a narrative account to make the statistical correlations historically meaningful. Among her findings are the discoveries that prior organization by labor aristocrats sometimes facilitated the Knights' recruitment of less skilled workers and that ethnic diversity did not impede unionization. Voss's data also support recent claims by social historians that community ties rather than workplace identities paved the way for broadly based labor organization.

The Knights of Labor reached formidable size in about two years and collapsed as quickly. Invoking political process approaches to social movements, Voss argues that external opposition doomed the Knights, not ethnic or occupational divisions within the working class. The union's meteoric rise spurred a counterattack by employers. With the state remaining neutral, employers won. Voss thus reaches by a different route the same conclusion as other recent studies of American exceptionalism: The actions and policies of employers and the state better account for historical outcomes than do the interests or culture of workers. Here, again, Voss draws on comparative evidence to show that employers in Britain and France were less combative or less powerful than their American counterparts (although she leaves unexplained U.S. employers' extraordinary antipathy to unions).

Voss concludes that "American industrial relations and labor politics are exceptional because in 1886 and 1887 employers won the class struggle" (p. 232). Rethinking the problem of exceptionalism as a battle between Knights and employers in 1887 makes it easier to explain the outcome. But can a defeat in 1887, however resounding, account for the exceptional character of the American labor movement from the late 19th century to the 1930s or later? Voss offers two main reasons for thinking so. First, defeating the Knights meant destroying the organizational base for working-class solidarity; this allowed sectional trade unionism (the AF of L) to gain strength and crowd out subsequent rivals. Second, the Knights' defeat discredited the vision of unity and a commonwealth of producers that had mobilized workers. These are reasonable arguments and commendable in highlighting the causal importance of historical sequence, timing, and legacies. They remain provisional hypotheses, however. Although Voss suggests that defeat bred resignation and a narrowing of vision among the next two generations of American workers, she does not document this "cognitive encumbrance." Nor does Voss make clear why AF of L unions, which experienced similar employer hostility and crushing defeats, would not, in turn, have been discredited and supplanted. Voss's account of the American labor movement's limited solidarity over the long haul is thus less convincing than her explanation for the Knights' rise and fall. But the latter is an impressive accomplishment: an exemplary application of social movement theory, statistical analysis, and historical sensibility to an important example of working-class mobilization.

*Regulating the Social: The Welfare State and Local Politics in Imperial Germany.* By George Steinmetz. Princeton, N.J.: Princeton University Press, 1993. Pp. xiv + 375. \$45.00.

Abram de Swaan

*Amsterdam School of Social Science Research*

Throughout Western Europe and in the United States the 19th century was a period of incisive reform of poor relief. Initially, the initiative was mostly with municipal governments that took over many of the tasks that villages and parishes could no longer deal with, but in the course of the century the central state became increasingly involved in the coordination and eventually the implementation of social policy.

In bringing together developments at these different levels as they occurred in Germany, George Steinmetz has written an exceptional book, revealing both the overriding continuities and the striking discontinuities between local and national policies. Moreover, the author set out to achieve "a merger of history and social theory" (p. 11), a task achieved more effectively in the course of his narrative, which is indeed thoroughly informed, even saturated, by historical sociology, than in his explicit discussion of general theoretical issues. This discussion deals in quick succession with debates that once appeared obligatory fare but now seem somewhat obsolete: the opposition between an explanation of social policy in terms of an "autonomous state" versus one in terms of "social factors" has in effect been superseded by the notion of a "relatively autonomous state" forced to deal with various societal actors, such as organized labor, landowners, entrepreneurs, and a panoply of reformers and professionals. This approach is precisely the one Steinmetz, rightly, adopts in his historical account of the rise of the Bismarckian welfare state. The antagonistic role of small entrepreneurs, shopkeepers, and artisans, first highlighted by Henri Hatzfeld (*Du paupérisme à la Sécurité Sociale: Essai sur les origines de la Sécurité Sociale en France, 1850–1940* [Paris: Colin, 1971]), remains somewhat underexposed in Steinmetz's account.

The author maintains (p. 215) that "the formation of social policy is a complex process that can rarely be explained in terms of a single causal mechanism or 'theory of the welfare state.'" Indeed, single-cause explanations hardly have a place in the social sciences, and the somewhat monolithic theories that the author presents in his review of the literature are indeed inadequate, but a more subtle and flexible theoretical construct does go a long way in explaining the main events in the book.

In a crucial passage (p. 144) Steinmetz boldly presents his own model of the "semiautonomous state" that "can explain Prussian and German social policy during the entire period from the 1830s to 1914. Both poor relief and worker policy can be accounted for in terms of the same causal factors. These factors take different forms in each period, however." First, there was "social fear," fear of fragmented social disorder in the

period of poor relief and of the organized socialist movement in the age of "worker policy." Second, there was the interventionist and effective German bureaucracy. Third, "heavy industry was the key element of the social coalition." Finally, four competing paradigms of social regulation both "shaped social policies and resulted from them."

In fact, a theory along these lines can explain a good part of the emergence of welfare states in other countries, also. Municipal, and in a later stage national, elites perceive the presence of the idle and laboring poor in their midst as both a threat (of crime, contagion, disorder, revolution) and an opportunity (a labor reserve, potential recruits, consumers, voters). Various paradigms to deal with this menace dominated the contemporary discourse. However, I would add, the elites faced dilemmas of collective action in dealing with the poor, since some in their ranks might profit from the effort without contributing to it. Coordination at the municipal level could solve this in an urban context. But the dilemmas did crop up again at the national level when cities were tempted to rid themselves of their disabled poor at the expense of other municipalities. Only compulsory, collective arrangements at the national level could overcome this predicament.

Steinmetz identifies the actors in the German constellation as an "activist" government in coalition with capitalist entrepreneurs, in this case preempting and excluding the workers' movement. However, elsewhere, organized workers were also part of the founding coalition of the welfare state, while the activist regime was the one indispensable partner everywhere. In other countries, too, the specific composition of the founding coalition and the oppositional strength of the small bourgeoisie and agrarian interests go far in determining both the timing and the nature of the resulting welfare state arrangements.

This book would have gained from a more comparative approach and from an analysis of the evolving system of states within which the German reich operated. But Steinmetz's careful and imaginative account of the rise of the German welfare state represents an instance of the cumulative character of historical sociology, greatly adding to the already considerable knowledge gathered so far and providing a mighty impetus to further analysis.

*Governing Capitalist Economies: Performance and Control of Economic Sectors.* Edited by J. Rogers Hollingsworth, Philippe C. Schmitter, and Wolfgang Streeck. New York: Oxford University Press, 1994. Pp. 316. \$45.00.

Peter A. Hall  
*Harvard University*

The study of comparative political economy, in sociology as in other disciplines, has long been dominated by a concern to show how institu-

tions with roots in wider social and political orders affect economic performance. Only a decade ago, a reader of this literature would have seen a stable set of theoretical frameworks built around the concepts of neocorporatism, strong or weak states, alternative financial systems, and "flexible specialization" versus mass production.

In the last decade, however, it has become apparent that the political economies of the industrialized nations are in the midst of a radical transformation associated with increasing internationalization, a technological revolution, and the reorganization of production. The institutional orders that once seemed to govern economic performance are in flux, and political economists have gone in search of new theoretical frameworks with which to conceptualize the contemporary political economy.

*Governing Capitalist Economies* constitutes an important contribution to this search. Based on a research group that met for the first time in 1988, it provides both testament to the trajectory of the field and empirical evidence on which further progress can be based. In an introduction and conclusion that are admirably careful and wide-ranging, the editors delineate a sweeping conceptual framework. They see the capitalist economy as an institutional order coordinated by specific governance systems defined as "the totality of institutional arrangements—including rules and rule-making agents—that regulate transactions inside and across the boundaries of an economic system" (p. 5). The principal object of the volume is to identify the governance systems of a variety of economic sectors in eight industrialized nations, to explain how these arise and change, and to explore how they affect firm behavior and sectoral performance.

The contributors move beyond the traditional distinction between markets and hierarchies to examine alternative governance systems based on informal networks of trust or subcontractors or associations formed from mutual interest or shared identity. Three overarching questions anchor the enquiry. How do governance systems affect economic performance? Is institutional variation by sector or by nation more important? To what degree are the governance systems of sectors converging in the face of increasing international pressures?

These are important and interesting questions. Although not every contributor focuses on all three of these issues, the volume sheds considerable light on them. Several tentative conclusions emerge. First, national differences still seem to have a major impact on sectoral behavior. Although there can be clear differences between sectors in the same nation, sectoral structure and behavior seem to vary substantially across nations. It is interesting that a group using the sector as its unit of analysis should find that national differences in institutions still matter. Second, we find such national differences because the state seems to have had an important impact on all the sectors studied here. More often than not, it is the legal order and material support provided by public authorities that allow specific associational orders to flourish. Third, despite some convergence in the most internationalized sectors, such as chemicals, consumer elec-

tronics, and securities, the contributors find little evidence of global convergence in the face of market pressure.

Most of the chapters of the book can be read for their own sake, as detailed studies of sectoral development in comparative perspective. Among the best of them are a finely-focused study of the circuit-board industry in Britain and Japan by M. Sako, an examination of the impact of the "industrial order" on machine tools in the United States and Germany by G. Herrigel, a study of the evolution of consumer electronics in Britain and France by A. Cawson, and a comparison of the dairy industry in Britain, Germany, and Austria by F. Traxler and B. Unger.

The volume also contains informative studies of the American and Japanese steel industries by P. O'Brien, the shipbuilding industry in Japan, Sweden, and Germany by B. Strath, the chemical industry in Germany and Britain by W. Grant and W. Paterson, the automobile industries of Germany, Britain, and France by B. Dankbaar, and the securities industry in Canada, the United States, and Britain by W. Coleman.

The reader may be somewhat disappointed to find that the book does not provide a clear typology categorizing sectoral governance across all of these cases or a systematic assessment of the impact of these governance systems on economic performance. However, it is not easy to do this kind of detailed sectoral research across nations. The book is a valuable exploratory exercise. As such, it takes us some distance forward by providing a wealth of hard-won data framed by an intelligent set of conceptual questions. There are real nuggets of wisdom in these chapters, which delineate well the institutional linkages between the social order and firm behavior. It is on studies such as these that the reorientation of the field will be based.

*Forging Industrial Policy: The United States, Britain, and France in the Railway Age.* By Frank Dobbin. New York: Cambridge University Press, 1994. Pp. xii + 262.

Richard Rubinson  
*Emory University*

Why do modern nations have such different political policies for organizing industries? *Forging Industrial Policy* is an elegant study designed to develop a cultural theory of these national industrial policies through a comparative study of the railroad industry in the United States, Britain, and France.

The study begins with the observation that each country's railroad policy has been constructed around a core conception that identifies both the conditions of industrial growth and the threats to industrial efficiency. United States policy was built around protecting the free market to ensure price competition as the mechanism of efficiency and preventing concen-

tration through either state power or industrial monopolization as the threats to that efficiency. French policy was built around the concentration of state power over the industry as the mechanism to ensure that private interests would serve the national good, and capitalist entrepreneurs and market irrationalities were conceived as the primary threats to that economic efficiency. British policy was built around protecting the small entrepreneurial firm as the source of industrial efficiency and identifying unbridled price competition and the predatory grab of large firms as threats to that goal. Thus, the United States gave authority to markets, France to the state bureaucracy, and Britain to the individual firm.

Dobbin explains these differences by developing a model of isomorphism derived from neoinstitutional theory that argues: (1) The root cause of national industrial policy rests in the distribution of political power, which imposed political order in the historical process of state formation. (2) These structures of state power and authority then led to conceptions of political order, a political culture, that symbolized this distribution of power. (3) These conceptions of political order were then "teleologically reinterpreted" (p. 21) as principles of economic order and authority that gave meaning to different possible industrial policies. And (4) these meanings shaped industrial policy by defining which policy solutions were perceived as efficient and moral and what actions were perceived as problems and destructive of economic efficiency.

The process by which this institutional isomorphism occurs, linking patterns of state formation to industrial policy, is not one of simple diffusion. Rather, political culture leads to industrial policy as disputes are worked out among the many different parties that contend over the large number of issues that arise in the construction of the railway industries. The political culture comes to forge an isomorphic industrial policy by giving clear rhetorical and logical authority, and consequently political advantage, to groups whose industrial policy solutions are isomorphic with that political culture. The historical analysis traces these disputes through the areas of railroad planning, financing, technical and managerial coordination, and pricing and competition. Most of the book is a detailed account for each country of the development of this isomorphism in each of these areas through the historical debates and their resolutions.

*Forging Industrial Policy* is a bold approach to developing a full cultural theory rather than just applying a cultural analysis to railroad policy. Dobbin argues that recent cultural approaches in sociology are far too narrow since their authors identify cultural institutions as only those that are "symbolic" and then are content to show how the symbolic realm affects the instrumental realm. In Dobbin's conception, like the Durkheimian tradition from which his work derives, all social institutions are cultural, in the fundamental sense that they are constructed systems of symbolic meaning, and so he seeks to demonstrate this approach in instrumental, "rational" institutions like industrial policy. Correspond-



ingly, he also argues that the three major social science theories used to explain industrial policy—interest group or class theory, statist theory, and economic theory—should be understood as symbolic meaning systems, or ritual myths, that social science employs to make sense of institutional patterns, rather than as sociologically valid explanations. That is, rather than explaining why class or statist or economic theory has developed as part of our society's meaning systems and so affects its policies, social scientists uncritically accept one or the other of these theories as valid explanations for policy. Although Dobbin never self-reflexively applies this same constructivist critique to the recent social science renewal of cultural theories such as his own, his book will still enrich sociological analyses by showing how a revitalized Durkheimian tradition can be applied to the areas of political and economic analysis, much as revitalized Marxian and Weberian analyses have already transformed our understanding of these areas.

*Manufacturing Militance: Workers' Movements in Brazil and South Africa, 1970–1985.* By Gay W. Seidman. Berkeley and Los Angeles: University of California Press, 1994. Pp. x + 361.

Kevin Neuhauser  
*University of Washington*

When I first heard about *Manufacturing Militance*, I must admit that I had my doubts. How useful could a comparison of labor movements in two such dissimilar countries be? After all, South Africa and Brazil represent diametrically opposed ways of constructing the intersection of race and class. In South Africa the state created a crude tripartite racial definition of the working class, each with different legal rights and restrictions. In Brazil the working class covers a subtle, culturally defined racial continuum that ostensibly enjoys equal legal standing but in practice experiences racial discrimination.

Sometimes, though, taking risks is justified, and this is one of those times. Seidman presents a case for convergence between the workers' movements of Brazil and South Africa that is all the more compelling precisely *because* of the obvious differences. Seidman argues that both movements represent instances of "social-movement unionism," in other words, "an effort to raise the living standards of the working class as a whole rather than to protect individually defined interests of union members" (p. 2). This convergence was produced by similar state policies that simultaneously created a demand for more skilled labor in the dynamic sectors of the economy, focused worker dissatisfaction on the state (because of its role in labor control), and forced those relatively skilled workers to share the deteriorating living conditions of poor urban communities. When economic crises weakened the capitalist-state alliance,

the stage was set for the rapid emergence of workers' movements demanding both economic and political restructuring of society.

Seidman's careful analysis is impressive, given the need to master two sets of literature that rarely if ever overlap. To someone who knows only one of those literatures (Brazil), the argument is as convincing as a two-country comparison can be. In fact, Seidman does a nice job in the conclusion of briefly drawing comparisons with South Korea (similar economic and political context) to reinforce her case, while at the same time identifying differences that may account for some variation in the workers' movements.

With such a small number of cases, however, one lingering doubt remains. Are the crucial common factors identified in Brazil and South Africa sufficient to produce social-movement unionism? One troubling case that comes to mind is Mexico—the Mexican state has pursued a similar industrialization program and has been actively engaged in controlling labor, and Mexican urbanization has followed a similar trajectory. Yet, the economic crisis of the 1980s did not produce "social-movement" labor mobilization in Mexico. This suggests that the factors identified by Seidman may be necessary but not sufficient to produce social-movement unionism. For example, what kind of state-capitalist relationship allows or forces a breakdown of the alliance under economic stress? Are there forms of state labor control that are successful in preventing mobilization on a large scale even when the other factors are present?

These questions are not meant as criticism of *Manufacturing Militance*, quite the contrary. They indicate the richness of an analysis that provokes new questions and the need to make new comparisons. Seidman introduces one theme in particular that merits serious research attention—the role of women workers as mediators between the workplace and the community. Seidman suggests that women may have been particularly important in connecting workplace issues to community issues because their reproductive and productive roles could not be kept separate. This pushed male coworkers to see the importance of community issues and forced community leaders to recognize the need to support workplace struggles. It was precisely this merging of workplace and community that made the workers' movements so powerful in Brazil and South Africa. Thus, women's significance may have been much greater than their numerical presence in unions and union leadership, because they were crucial actors in creating "class-conscious" labor movements.

In sum, *Manufacturing Militance* is an example of the power of carefully chosen comparative studies. A two-country study cannot answer all questions about the emergence of social-movement unionism and Seidman never makes this claim. But, after reading the book, it is hard to argue against the conclusion that structural similarities in the political and economic forms of dependent development pursued by Brazil and

South Africa pushed labor mobilization in common directions in the two countries—despite the many cultural and historical differences.

*Looking for God in Brazil: The Progressive Catholic Church in Urban Brazil's Religious Arena.* By John Burdick. Berkeley and Los Angeles: University of California Press, 1993. Pp. xii + 280.

David A. Smilde  
*University of Chicago*

In *Looking for God in Brazil*, John Burdick seeks to explain two "paradoxes" surrounding the "Christian base communities" (CEBs) of the progressive Catholic Church. To do so, he undertakes an ethnographic study of participation in CEBs, Pentecostalism, and *umbanda* (a rapidly growing Afro-Brazilian religion), in a *bairro* (a town within a municipality) north of Rio de Janeiro.

He first addresses the "numerical paradox," asking why the membership rolls of the CEBs are undramatic when compared to the growth of other religious alternatives (p. 5)? The core of his answer comes in a distinction between religions of "continuity" and "discontinuity." The CEBs, behind the image of Christians living to reshape the world, "remain aloof from notions of radical rupture in the self, emphasizing instead continuity between religious and nonreligious roles and statuses" (p. 223). Pentecostalism and *umbanda*, on the other hand, "as cults of discontinuity or transformation, conceive of radical ruptures in the self as both possible and desirable" (p. 224).

Using this basic distinction, Burdick sets out to show how those in the lower class, women, youth, and racial minorities cope "through available religious acts and language, with their experiential predicaments" (p. 9). First, various aspects of the liberationist project unwittingly highlight the social hierarchy of the *bairro*. For example, the practice of meeting in homes of participants and the emphasis on literacy are humiliating to those who have less adequate material surroundings or do not read well. The Assembly of God, on the other hand, by emphasizing separation from the "world," denouncing vanity, and valuing the "spirit" over the "word," diminishes these painful class distinctions.

Second, women find that the CEB emphasis on political action means that the domestic issues that most immediately affect them are considered inappropriate for discussion groups. *Umbanda* and Pentecostalism, in contrast, provide them with a sympathetic atmosphere removed from the larger social arena (and thus gossip) in which blame is placed on "outside humanlike forces." Pentecostal ideology offers female subordination in exchange for domestic peace but provides a language in which models of gender are contested.

Third, for adolescents who wish to escape youth status culture, the

CEBs do not present an effective alternative. "In the absence of a transformative discourse, youthful rivalries, jealousies, and status rankings from school and street are transferred intact to the religious setting" (p. 127). The Assembly of God, on the other hand, explicitly rejects worldly status competition and provides youth with concrete social alternatives.

Fourth, the CEBs do not provide an effective counterdiscourse to racism. The emphasis on continuity results in an unintended continuity with the racism of the larger social world. Both *umbanda* and Pentecostalism, in contrast, through their emphasis on personal charisma and the spirit, offer a positive spiritual identity and access to positions of authority.

While the author analyzes problems of class, gender, adolescence, and race, he makes curiously little reference to problems of economic subsistence. The omission is conspicuous not only because most Latin Americans encounter economic subsistence as an all-consuming "experiential predicament" but because economic "adaptation" has been one of the primary explanations for movement to evangelical groups (see David Martin, *Tongues of Fire* [London: Basil Blackwell, 1990], chap. 11, for a review of the literature). While including economic adaptation among his explanations would have strengthened his case, it also would have required a more complex analysis of the circuitous cultural logic of "otherworldly" religion than his continuity/discontinuity distinction provides.

The last chapter addresses the "political paradox": the presence of "less than socially activist tendencies" among many CEB members. Through a clever study of the reception of the liberationist message, Burdick shows that actual penetration has been "uneven and uneasy." He finds Pentecostals, on the other hand, to be less otherworldly than reputed. Because of its radical egalitarianism, Pentecostalism "carries as much long-term potential for becoming a religion of revolution as does liberationist catholicism" (p. 226).

The principal shortcoming of the book lies in its inadequate ties to existing literature on the issues examined. For example, in the last chapter Burdick makes no reference to work of Ernst Troeltsch or Max Weber, which the section largely corroborates and which could have provided theoretical backbone. Throughout the book, the author repeatedly speaks of social movements while mentioning only a few obscure works in the field. Had he touched on the principal debates, his "paradoxes" would instead have looked like the classic problems of social movements. In addition, his introductory discussion on theory and method does not reflect a familiarity with the literature on religious choice and conversion and presents as a theoretical advance a "model" that is little more than a call for proper scientific method.

Still, what *Looking for God in Brazil* does far overshadows what it does not do. It provides the literature on Latin American popular religion with a lucid, coherent, and long-overdue ground-level ethnography of class, gender, and politics and convincingly brings the factors of youth and race into the debate.

*Religious Melancholy and Protestant Experience in America.* By Julius H. Rubin. New York: Oxford University Press, 1994. Pp. ix + 308. \$35.00.

Richard W. Flory  
*University of Chicago*

Julius H. Rubin has written a compelling account of what he believes to be a uniquely Protestant experience, the quest for individual salvation that ironically results in the believer oscillating between the experience of euphoria in the presence of God, followed by a sense of despair at his abandonment. Rubin argues that this has been a defining feature of Protestantism since the Reformation, when the Catholic emphasis on forgiveness of sins through confession was rejected, leaving the promises of "God's assurance of election and salvation" outside the grasp of the believer (p. 3). He embeds his analysis in the context of Weber's "Protestant ethic" thesis, that just as practical rationalism and the regulation of life conduct led to world mastery and ultimately the growth of modern capitalism, it also led believers to pursue what might be termed the mastery of the self over sin, often resulting in "distinct forms of melancholy and mental disorders" (p. 21).

Rubin's focus is on "religious melancholy" as a distinctive mental disorder, which he defines as a crisis of "spiritual passage and conversion in which each penitent felt forsaken by God's love" and further, that this was a "distinctive psychopathology characteristic of evangelical Protestants" (p. 5). Using diaries and spiritual biographies of 18th- and 19th-century Protestants and case histories of patients at the Hartford Retreat, an asylum for the mentally ill, Rubin details the various manifestations of religious melancholy in the lives of evangelical Protestant believers. These individuals, seeking assurance of salvation and cleansing from sin, entered into various forms of self-induced torment, such as excessive grieving over personal adversity that they interpreted as God's abandonment and lengthy periods of fasting ("evangelical anorexia nervosa"), meant to cleanse the believer from the bout of melancholy and sin, but which often resulted in death.

With the rise of the asylum movement and the medicalization of various mental disorders, the diagnosis of religious melancholy moved from the realm of pastoral care to the asylum, and the spiritual crisis experienced by patients was defined as epiphenomenal to the underlying mental illness, thus obscuring the real roots of their affliction. Rubin maintains that religious melancholy disappeared in the late 19th-century as Protestant modernism replaced evangelical Pietism, only to reappear in the 1950s with the rise of "new evangelical" revivalists such as Billy Graham and their emphasis on individual salvation and the attendant doubt and depression that accompanied the quest for religious assurance.

The spiritual biographies that Rubin details are fascinating and provide insight into the personal battles of a distinct group of Protestant

believers. However, his insistence on the central role of evangelical Protestantism in the incidence of religious melancholy is not entirely convincing and raises more questions than it answers.

Rubin states early on that it is impossible to know the true incidence of religious melancholy, yet he argues that the case histories and spiritual biographies of the "religious melancholiacs" that he presents are representative of the experience of evangelical Protestants. The fact that his cases were drawn from a distinct subgroup of disturbed believers leads one to suspect the generalizability of his findings. While it is certain that many Protestants exhibit moments of personal doubt, it does not necessarily follow that extended periods of religious melancholy will result nor that this is a core feature of the Protestant religious experience as exhibited by most believers.

A related issue is the nature of the relationship between evangelical Protestantism and religious melancholy. Throughout most of the book, Rubin argues that it is the Protestant "morphology of conversion" that forces the individual to continually search for the assurance of his or her salvation, resulting in religious melancholy. By arguing for what is in effect a causal relationship, Rubin never seriously considers the possibility that other factors, such as gender, class, or family background, might correlate with the descent into religious melancholy. Was it more prevalent among women than men or among the middle and upper classes? Could the poor afford to be afflicted by religious melancholy? These questions remain unanswered.

By the end of the book, however, Rubin does reveal the possibility that at least for 20th-century sufferers of religious melancholy, feeling forsaken by God was a result of personal adversity: "The crisis engendered by feeling forsaken of God involved the perception of abandonment by Him, usually caused by personal adversity" (p. 237). This change in emphasis is not explained or even noted, but it raises a question that must be pursued: Why do Protestants in the late 20th century experience religious melancholy because of personal adversity, while those in the 18th and 19th centuries encountered religious melancholy as a result of the intense religious quest for salvation?

This book charts new territory in the sociological study of religion and presents a thesis that as it is debated and refined will yield new insights into the relationship between evangelical Protestantism and psychological phenomena. Moreover, this book could provide the social and historical background for an analysis of what I see as a related 20th-century development in evangelical Protestantism, the professionalization of a distinctly evangelical Christian psychology that explicitly seeks to integrate evangelical teaching with psychoanalytic theory. This in itself would be a worthy contribution.

*Practicing Virtues: Moral Traditions at Quaker and Military Boarding Schools.* By Kim Hays. Berkeley and Los Angeles: University of California Press, 1994. Pp. x + 289. \$30.00.

Fred Kniss

*Loyola University of Chicago*

*Practicing Virtues* is a refreshingly good read. Hays explores an inherently interesting topic, using an incisive analytical model and lucid prose to present a clear set of arguments.

Contrary to the impression that might be given by the title, this book is not primarily a contribution to the sociology of education. The schools Hays studied are clearly not typical educational institutions. Instead, Hays uses her fieldwork in three Quaker high schools and three military high schools to provide an empirical demonstration of Ann Swidler's notion of "culture as a tool kit." She addresses the more general question: How are the core virtues of particular moral traditions used in the everyday lives of people who espouse them and in the institutions that embody them? She suggests that Quaker schools represent the American liberal tradition while military schools typify American conservatism. Given the American book-buying public's current interest in "culture wars" and issues of public and private virtue, *Practicing Virtues* has the good fortune to be both sociologically significant and timely.

Hays draws a sharp contrast between the Quaker and military moral traditions. She characterizes the Quaker tradition as typically liberal because it sees the source of morality as internal (the Quaker notion of a godly "Inner Light") and sees the primary target of moral action as the community, its purpose to transform social structures into ones that embody core virtues and facilitate the nurturing of the Inner Light. For Quakers, the core virtues are equality, community, simplicity, and peace. She characterizes the military tradition as typically conservative because it sees the source of morality as external (military standards and discipline) and sees the primary moral project as the transformation of individuals, producing persons whose behavior reflects the core military virtues of loyalty, competence, selflessness, integrity, and pride.

Hays then shows how the core virtues of these contrasting moral traditions have contrasting implications for people and institutions. On one hand, they provide order as they are instantiated in particular structures and institutionalized practices in the schools. For example, in the Quaker schools, programs, curricula, and decision-making practices are organized to nurture the unique Inner Light of each individual student. This produces a focus on teacher-student relationships and individualized attention to educational evaluation and behavioral issues. By contrast, the military schools' goal of producing disciplined, competent individuals who are loyal to external standards produces distance between teachers and students and an emphasis on "maintaining standards" in academics and extracurricular behavior.

In addition to providing order, however, moral traditions are also sources of conflict, as differing interpretations of core virtues clash when they are practiced in real life situations where individual and communal interests may not be identical. Hays proposes a set of conflict types that are faced by both Quaker and military schools. Within this typology, the contrasting core virtues act as lenses and filters that highlight certain kinds of issues and exclude others. For example, one conflict type is the conflict between general principles and individual needs. This type creates a problem for Quaker schools when they must decide how to treat the Inner Light of politically conservative students who espouse militarism. By contrast, the military virtue of competence leads to conflict over grading between those teachers who uphold the general principle of high academic standards and those who are willing to make accommodations for particular individuals with particular needs. Hays's central argument here is that such conflicts serve to strengthen the moral tradition over the long run, because conflicts reaffirm individuals' commitments to the virtues and adapt interpretations of the virtues to new situations.

In other chapters, Hays explores how the contrasting traditions produce different decision-making procedures and different ways of exercising authority within the schools. In the final chapter, she examines how the moral traditions interact with the special developmental tasks faced by adolescents. This chapter is not well integrated with the rest of the book, but I found it to be the most interesting of all. She shows how adolescents adopt and adapt elements of the traditions, infusing them with special meaning that is relevant to adolescent struggles. What is missing in the adolescent appropriation of the virtues is the identification of them with the "sacred." The schools' programs and faculty intend to teach that practicing virtue is a "sacred" activity; that is, it involves the absorption of one's self into something larger that transcends the self. But, Hays argues, this task is actually antithetical to the central task of adolescence—identity construction and differentiating one's self from others. Thus, adolescents appropriate the moral tradition partially and selectively. Still, Hays suggests that the model the schools provide of the sacred character of the virtues is available for full appropriation by students later in life after their individual identities are more fully developed.

It is probably clear by now that Hays's analysis is explicitly and unabashedly Durkheimian and that it shares Durkheim's optimistic trust in the generally benign character of moral orders. Many readers, myself included, might wish for a bit more Weber or Marx in the account. That is, Hays offers little insight into how particular interpretations of virtues and the practices that result might be tied to the organizational, political, or material interests of those who espouse them. Although there are occasional references to the potential "dark side" of virtuous practices, we are not given a very clear picture of what happens to the losers of contests over interpretations of virtues or to those who are excluded by a particular moral order. But these are quibbles from a reader who does not share



the author's Durkheimian worldview. They should not obscure the fact that this is a fine piece of research that is also a lot of fun to read.

*Inside Culture: Art and Class in the American Home.* By David Halle. Chicago: University of Chicago Press, 1993. Pp. xvi + 261. \$29.95.

Eugene Halton  
*University of Notre Dame*

If the expression "seeing is believing" can be taken literally, then take a good look around your home, especially at your wall hangings. Do you have landscapes, abstract art, primitive art, photographs, or religious icons? What do these things say about you and your beliefs? Are you a "cultural capitalist" theorist by day, writing tirelessly of the dominant and dominated classes, who yet by night relaxes amidst abstract art pieces whose very presence indicts you as a member of the dominating elite? Or have you established props of the working-class hero drawn from popular culture iconography? Or do your visual representations speak meanings other than class or education? Are they symbols of family ties, or perhaps signs of the distance you have made from your family? Are your landscapes populated or depopulated? Does what you have on your walls or mantels relate to whether you live in the city, suburb, or country?

These kinds of questions are systematically dealt with in David Halle's *Inside Culture*, which surveys the visual representations displayed in American homes. By turning the focus of art from the artist who made it, the museum that displays it, and the critic who analyzes it to the home and the people who live with it, a very different perspective on art, class, and culture emerges.

The study is based on a sample of 160 houses from four areas in and around New York City, composed of upper- and upper-middle-class Manhattan residents, upper-middle-class suburbanites, and working-class residents from the city (Greenpoint) and suburbs (Medford), as well as subsamples including exclusive vacation houses in the Hamptons and a middle-class black suburb (Spinney Hill). Halle's first chapter, on the house and its context, provides an interesting portrait of the architectural, historical, ethnic, and racial backgrounds of the neighborhoods and houses studied, which helps to explain a number of patterns found in the objects within the homes. The numerous photographs throughout the book not only convey the data discussed, but provide good food for sociological contemplation.

Halle amply demonstrates how the typical American home is an ongoing show-and-tell of the beliefs of its inhabitants, revealing the myriad influences—conscious and unconscious, conventional and personal—that make up the meanings of art and the self. Social class has real consequences for the selection and valuation of artistic, photographic, and

religious images in the home. Yet, according to Halle, the relations between the meanings people attribute to their art and dominant classes and artistic elites, "is poorly understood if viewed exclusively through the lens of domination and power" (p. 204). Halle is not saying bye-bye Bourdieu, farewell Frankfurt, and so-long to status-striving theories of culture so much as insisting that culture is more fluid and multi-determined than these theories usually admit.

Take landscape depictions, for example, which were the most common type of object in the homes sampled, making up about one-third of all objects for all classes. Landscapes seem to function as a visual background opiate of the people, regardless of class. Only two of the 349 landscapes did not depict calm scenes, and those two have their own stories to tell.

Contemporary American landscapes are overwhelmingly depopulated, while people appear more frequently in historical or foreign landscapes. These contemporary American landscapes represent an ideal of a calm, depopulated landscape for all classes, reflecting, in Halle's interpretation, the suburbanization of life—including urban life—as well as the privatization of the domestic sphere, with its decline of extended family members and boarders and transformation of the backyard from the toilet zone—the outhouse—in the early 20th century into a leisure zone in more recent decades.

As the contemporary landscape representation became depopulated, homes seem to have become increasingly peopled by family photos. Yet the photos speak not only of family connectedness, but of the fragility of the contemporary family and protean course of the self. Halle's interpretation of the racial elements and political dimensions of primitive art, though contestable in my opinion, offer an unexpected angle on segregation and the encounter with the "other." The book is well worth reading, especially in your own home.

*Claims to Fame: Celebrity in Contemporary America.* By Joshua Gamson. Berkeley and Los Angeles: University of California Press, 1994. Pp. xii + 256. \$35.00 (cloth); \$14.00 (paper).

Michèle Lamont  
*Princeton University*

In *Claims to Fame*, Joshua Gamson sets out to examine the nature of entertainment celebrity in America, focusing on its meaning, its production, and its consumption. The author mobilizes a panoply of techniques to probe the full scope of the celebrity phenomenon, analyzing the essence of fame as defined by *TV Guide*, interviewing Los Angeles agents, performers, publicists, and journalists, and observing Arsenio Hall's fans to determine what makes them tick. The result is a nicely written, insightful, informative, and multifaceted book, an enjoyable volume that

will fascinate those interested in the changing cultural makeup of American society.

The author believes that the understudied celebrity culture deserves analysis because it occupies "a large space in many Americans' daily lives" (p. 6) and embodies commercial culture at work. He approaches this culture both institutionally and interpretatively by zooming in on the organizational interdependencies that constrain people's work and on the negotiating processes they generate. He also probes the filters through which celebrity producers and consumers experience celebrity. He states his case plainly: in the celebrity world, things do not "just happen" (p. 69). This world is an industry like many others, with division of labor, rules, conventions, and career paths. What really interests Gamson is the connection between discourse, production, and reception, that is, "How do production processes and interest-driven activities affect what is said and received? . . . How do audiences understand production activities? . . . How do producers perceive audiences and build them into their own work?" (pp. 5-6). While addressing these questions, the author unpacks an unexplored terrain, enriching our understanding of celebrity by "mining the superficiality of this world for its depth" (p. 6).

In part 1, Gamson draws on magazines such as *People* and *Vanity Fair* to document how the discourse on celebrity has changed during the 20th century. He argues that the last 60 years witnessed a major change in fame technology as "fame as it rises from greatness" was gradually replaced by "fame as artificial production" (p. 16). He traces the history of an emerging industry of celebrity creation and discursive shifts in which money and resources, instead of talent, are increasingly viewed as determinants of success. At the same time, the meaning of fame itself changes as "famous people become entertainment instead of entertainment producing fame." It is in the fifties that celebrity creation becomes truly a business in itself. In part 2, Gamson draws on interviews with "celebrity producers" to discover how they understand the business of selling celebrity and how they define marketing strategies for new and established stars. In the picture that emerges, the artist is little more than clay to be molded, a shallow character for whom "art" is a minor consideration. Finally, in part 3, Gamson draws on focus-group interviews with celebrity-watching people to ponder "What makes people so fascinated with unreality?" and "How does unreality become one of our modes of reference?" The answer is, in part, the pleasure individuals get from sharing with others a world without consequence where familiarity does not involve responsibility.

Judging from the appendix, the author has not interviewed highly respected entertainers, only minor ones. I leave the book feeling that I have learned more about the production of Vanna White and Angelyne than about that of Meryl Streep, Anthony Hopkins, or Daniel Day Lewis. How do major celebrities live with themselves and think about their talent? Is money the gauge of their success? When asked whom they respect, do they point to Madonna, Sylvester Stallone, or Jeremy Irons?

Is there any notion of artistic sacredness left in their work lives? And to what extent is the celebrity industry an international phenomenon? Gamson does not address these questions. Does the type of work that entertainers engage in seriously affect the type of fame they produce? In other words, should not the content of the work itself be given more consideration if only as counterevidence to the industry-driven model proposed by Gamson? Furthermore, does the fame technology used in entertainment differ from that used in other realms? How does Vanna White compare with a Henry Louis Gates, Mother Theresa, or Nancy Kerrigan? Comparisons on these points would have been fascinating. Finally, do secretaries indulge in celebrity watching differently from how professionals do? The participants in the focus-group interviews are described as "middle class," but their occupations suggest that they constitute a less homogeneous group. It would have been interesting to analyze fan cultures with an eye for intergroup differences.

Gamson concludes the book by discussing the dangers of marketing Bill Clinton with Oprah Winfrey technology while appealing to Americans who want to view success as available to all by way of money. This raises broader issues about the trivialization of achievement and the triumph of money as the gauge of success in American society. Therein lies the true importance of this book.

*Nightwork: Sexuality, Pleasure, and Corporate Masculinity in a Tokyo Hostess Club.* By Anne Allison. Chicago: University of Chicago Press, 1994. Pp. xiii + 213. \$37.00 (cloth); \$14.95 (paper).

Yuko Ogasawara  
*University of Chicago*

This book is a welcome addition by an anthropologist to the study of aspects of Japanese culture that some may consider marginal. It introduces the readers to a dimension of Japanese white-collar workers' lives seldom considered seriously even by Japanese scholars themselves. Although research on an atypical topic is sometimes suspected of compromising quality for rarity, no such skepticism is warranted in the case of the present book. Allison beautifully combines Marxist and Lacanian theory to explain going to hostess clubs as a practice that suits the needs of corporations.

The principle of company recreation, exemplified by going to hostess clubs, is to entertain employees and clients at a place away from work in order to strengthen office or business relationships. According to Allison, medium-to-large Japanese companies spend as much as 5% of their annual operating expenses on corporate entertainment. The biggest spenders are large, prestigious companies, and those who frequent hostess clubs tend to be elite businessmen.

The primary data for the study comes from a four-month participant

observation in one of the top-drawer hostess clubs in Tokyo's fashionable Roppongi district. In the club, referred to as "Bijo," the author worked as a hostess, performing various ritualized tasks: lighting customers' cigarettes; pouring their drinks; flattering them with constant, exaggerated compliments; and bringing a group of businessmen at the table into an exchange by orchestrating a lively discussion and by encouraging everyone to sing. The service provided by a hostess, in other words, is essentially at the level of conversation facilitator.

One of the key components of the talk within hostess clubs is reference to the hostess's body parts, often her breasts. Comments such as "You have large breasts" or "Your breasts are nonexistent" are repeatedly made, always followed by laughter. The language and behavior at the tables are thus sexually charged, yet sex is confined to conversation. Allison argues that the anatomical talk is, in fact, not about a heterosexual relationship with a woman but rather a homosocial statement about being a man. By reducing a hostess to the topic of breasts, men turn her into an object about which they can come to agreement with one another.

It is Allison's contention that sexuality that is all talk not only structures the identity of women but that of men: When talking about a hostess's breasts, "Men assume an identity far more unifying and uniform than the one they are assigned through work during the day. Ranks of difference and possible discord are dissolved when all men become *sukebei* (desiring, lecherous males) in a club" (p. 181). Nighttime entertainment paid by corporations homogenizes male workers. Through a playful sexuality that titillates men's egos and makes them feel self-assured, a particular construct of "maleness" is developed that serves to align men with the workplace. Going to hostess clubs, therefore, is not primarily for or about sex. It concerns, rather, the subjectivization of workers.

One problem sociologists might find with the work is the casual way in which the author treats the information obtained from interviews. Not only do we not know how many persons were interviewed altogether, but attributes of interviewees that seem relevant, such as age, occupation, and the type of workplace (indeed, whether or not the interviewee is a white-collar worker in a large company who is likely to visit hostess clubs on company expense accounts) are not often specified. For example, little is disclosed about one man whose attitude toward a nightlife woman's value is introduced except that he was a student of an English class that the author taught (p. 174).

I should also point out that there are quite a few incorrect Japanese expressions, including terms widely used among Japanese white-collar businessmen. For instance, *hirashain* (rank-and-file employees) and *madoiwazoku* ("deadwood" employees) appear inaccurately as *heishain* (p. 89) and *madozaizoku* (p. 98), errors apparently due to a misreading of their *kanji* characters. It is most unfortunate that a book written with such perception and force should contain such easily avoidable errors as these, because, although they are themselves minor mistakes, they lead one to wonder about the author's literacy of the culture she is studying.

These problems, however, do not diminish significant contributions Allison makes to the study of gender identities and of Japanese corporate culture. The book will appeal to sociologists of culture, organizations, and gender and sex, as well as anyone interested in life in contemporary Japan.

*Femininities, Masculinities, Sexualities: Freud and Beyond.* By Nancy J. Chodorow. Lexington: University Press of Kentucky, 1994. Pp. xi + 132. \$20.00 (cloth); \$9.95 (paper).

Patricia Ticineto Clough  
Queens College, City University of New York

In *The Reproduction of Mothering* (Berkeley: University of California Press, 1978) Nancy Chodorow made clear that there is "no one to one correspondence between unconscious processes or structure and the content of consciousness and intended activity," that "a particular unconscious process . . . can express itself in almost endless behavioral as well as conscious psychological modes" (p. 41). Nonetheless, throughout the early 1980s, feminist critics used Chodorow's description of the intrapsychic effects of the pre-oedipal relationship of mother and daughter to argue for the common quality of women's experience, women's identity, even women's way of knowing.

Now, in *Femininities, Masculinities, Sexualities*, Chodorow redeems her earlier psychoanalytic insights about the discontinuity between unconscious processes and behavior or intended activity, while at the same time joining ongoing debates about subject identity and experience in relationship to race, class, ethnicity, nationalism, and sexuality. Chodorow especially focuses on heterosexism, which has concerned queer theorists as well. Indeed, like most queer theorists, Chodorow initiates her criticism of heterosexism by returning to Freud's thought.

In the first of the three chapters of *Femininities, Masculinities, Sexualities*, Chodorow rethinks Freud's writings on women and gives some sense of why and how Freud still matters. Of course, appreciating Freud involves redressing the accusations made against his thinking, especially his thinking about women. Chodorow does this by contrasting Freud's take on female sexuality, in which he assumes that there is only one kind of woman—perhaps best understood as a projection of a male's psyche—to Freud's "emphatically plural account of a multitude of 'women'" (p. 2), in which the subjectivities of women are more fully appreciated and their sexual desire more clearly recognized to be complex.

For Chodorow, Freud's recognition of the complexity of female sexuality depends on his (sometime) recognition of the difference between gender and sexuality—that is to say, that sexual identity, mode of desire, and object choice are not inherently linked to gender; indeed, they are never fixed except in defensive structurings. Queer theorists also have

emphasized that sexuality and gender are not reducible to each other or, at least, that there is a regulatory fiction that functions to converge sexuality, gender, and desire. For them, the deconstruction of this fiction not only implies a revision of Freud's thought but of feminist theorizing as well, since, for the most part, feminist theorists have treated gender as a social construction of sexuality, as if sexuality itself is not socially constructed as such.

For Chodorow too, rethinking Freud not only opens up to an analysis of gender that does not assume heterosexuality but therefore also opens up to a reconsideration of heterosexuality as a "defensive compromise formation," symptomatic of which is the taken-for-granted intertwining with heterosexuality of submission and male dominance. Chodorow thereby concludes that there are heterosexualities, homosexualities, masculinities, and femininities and that there needs to be further exploration of normalities and abnormalities that does not disavow these pluralities.

While all of this seems to contradict what cultural feminists made of *The Reproduction of Mothering*, Chodorow has no comments on this subject. She only reiterates what she first argued in *The Reproduction of Mothering*—that it is in relation to early experiences with their mothers that children shape their fantasies, identities, and modes of desire. But now Chodorow insists that "there is no single way" that these early experiences are worked out intrapsychically so as to distinguish all men from all women (p. 85).

Surely all of this leaves readers wondering: What now is to be thought about the relationship of unconscious processes and sociality? How is the relationship of oedipus, male domination, the law of the father, and the symbolic order to be treated in social and cultural criticism, especially feminist criticism? That Chodorow resists exploring issues of criticism is surprising and disappointing. That she makes no effort to address sociologists, as once she did, also is disappointing but perhaps less surprising; at least one can imagine that Chodorow has grown weary with most sociologists' reluctance to take seriously psychoanalysis or feminism but especially both together. We can only hope that she will return her efforts to engaging sociologists directly; sociologists need her ideas to help them rethink the relationship of unconscious processes and sociality.

*Fallen Women, Problem Girls: Unmarried Mothers and the Professionalization of Social Work, 1890–1945.* By Regina G. Kunzel. New Haven, Conn.: Yale University Press, 1993. Pp. 264. \$27.50.

Kristin Kay Barker  
*Linfield College*

This book is about the struggle over the meaning of "out-of-wedlock" pregnancy in the United States between 1890 and 1945. Kunzel skillfully documents how its meaning changed across time as different players in

this history—evangelical women reformers, social workers, psychiatrists, and single pregnant women—struggled to secure their interests. She provides a nuanced account of how the categories of gender, race, social class, and sexuality were tools employed in (and hence categories constituted in part by) the political battle over the institutional and cultural meaning of illegitimacy and professionalism.

Kunzel maintains that the history of dealing with illegitimacy between 1890 and 1945 is best understood as a tension between the moral reform efforts of evangelical women and the professionalization efforts of social workers. In the late 19th and early 20th century, evangelical women combined moral and religious purpose with the ideology of female difference to carve out their careers within the maternity home movement. These reform women established maternity homes for the redemption of their “fallen sisters”—victims of male seduction and abandonment. The melodramatic script of the “old, old story” (pp. 18–19) of cunning male sexuality and the innocent, trusting female justified the reformers’ benevolent care of single pregnant women by placing them outside of immorality (p. 20).

Like Sheila Rothman (*Women’s Proper Place* [New York: Basic, 1978]), Nancy Cott (*The Grounding of Modern Feminism* [New Haven, Conn.: Yale University Press, 1987]), and other social historians, Kunzel illustrates how women reformers of this period strategically used (and reproduced) the ideology of female difference to justify their own public sphere activities. Employing this strategy, evangelical women built an extensive maternity home network by the first decade of the 20th century. Their cultural authority over this domain, however, was challenged by the emerging profession of social work. Like other emerging professionals, social workers in the late 19th and early 20th century increased their educational and training standards, imposed standardized examinations, and founded a national professional organization. In spite of these hallmark indicators of professionalization, social workers’ professionalization struggles were hindered by the field’s female domination and, relatedly, its status as an “auxiliary” profession. Throughout the 1940s, the struggle for autonomy was a continuing battle for social workers generally and those working with out-of-wedlock pregnancy specifically (pp. 38–39).

In their professionalization struggles, social workers distanced themselves from the explicitly gendered language of evangelical reform in favor of what they perceived to be the gender-neutral language of professionalism. Social workers injected science into the discourse about illegitimacy and presented themselves as scientific experts. Casework diagnostics were part and parcel of this strategy. The diagnostic categories they employed not only carried the weight of science, they also recast unwed pregnant women. Whereas the evangelical women’s script of seduction and abandonment defined women as “fallen sisters” in need of redemption, social workers’ diagnostic categories of “feeble-mindedness” and “sex delinquency” defined these women as “problem girls” in need of scientific treatment (p. 37).



With the rise of medical psychiatry in the 1940s, social workers began to apply the psychiatric category "neurotic" to unwed mothers. While this category assisted social workers in further distancing themselves from reform women, it moved the discourse into the male-dominated medical jurisdiction, where women social workers held a marginal status. Ultimately, these contradictions within social work's professionalization strategy weakened their tenuous hold on illegitimacy. In their attempt to transform a female-dominated occupation into a profession, social workers "participated in the process of gendering professionalism in a way that equated professionalism with masculinity" (p. 48). Male psychiatrists replaced evangelical women and social workers as the cultural authorities concerning out-of-wedlock pregnancy.

The strongest aspect of Kunzel's argument is her specification of institutional changes that enabled social workers to gain access to the domain of illegitimacy. Unlike some historical accounts of how scientific claims come to garner cultural authority, Kunzel clearly documents the actual mechanisms that created an environment favorable to social work's scientific claims. The reader is not left wondering exactly how or why scientific claims gained the upper hand.

The weakest aspect of this book concerns Kunzel's portrayal of the role of single pregnant women in this history. While it is indeed likely that women did not utilize maternity homes in accordance with the agendas of evangelical women or social workers, Kunzel overly romanticizes women's resistance and agency. This result comes about primarily because she attempts to specify the role of single pregnant women based on historical documents produced by evangelical women and social workers. Kunzel is by no means unaware of this methodological dilemma, as she addresses it in some detail early in the text. Still, in her attempt to pull single pregnant women's voices out of her documents, she makes claims beyond those her data support. Kunzel's conclusions about the other key players in this history, however, are cleverly argued and logically based on her primary documents.

Kunzel's request to incorporate gender into our theoretical understandings of professionalization is well taken. Paradoxically, however, she did not fully utilize scholarship that has attempted to specify the ways in which professionalization has a gender bifurcation. Specifically, the accounts of nursing's professionalization attempts (Barbara Melosh, *The Physician's Hand* [Philadelphia: Temple University Press, 1982]; Susan Reverby, *Ordered to Care* [Cambridge: Cambridge University Press, 1987]), while cited, were not engaged in a comparative theoretical manner. Additionally, her failure to address several pivotal scholars makes her critique of the sociological work on professionalization problematic. Particularly troubling is her failure to incorporate Andrew Abbott's (*The System of Professions* [Chicago: University of Chicago Press, 1988]) concept of the "system of professions." Abbott's framework could have provided theoretical guidance to Kunzel's strong empirical project. Still, Kunzel's work makes an important contribution to the collective project

of building a more specified theoretical understanding of professionalization that highlights its gendered nature. For that, she is to be strongly commended.

The focus of this book will make it of general interest to scholars of work, professionalization, and medicalization. Additionally, those interested in the history of social work or of single mothers and social policy in the United States will want to read this book.

*Birth Weight and Economic Growth: Women's Living Standards in the Industrializing West.* By W. Peter Ward. Chicago: University of Chicago Press, 1993. Pp. xv + 218. \$38.00.

Jeremy Atack  
Vanderbilt University

Using records from six maternity hospitals in five major European and North American cities—Boston, Dublin, Edinburgh, Montreal and Vienna—W. Peter Ward, professor of history at the University of British Columbia, has written what is likely to be the definitive work on the historical record of birth weights during industrialization in the Western world. This study is an important supplement to the anthropometric work of Robert W. Fogel and his collaborators (much of which is summarized in Fogel's "Nutrition and the Decline in Mortality since 1700: Some Preliminary Findings" in *Long Term Factors in American Economic Growth*, edited by Stanley L. Engerman and Robert E. Gallman [Chicago: University of Chicago Press, 1986]). The fundamental proposition underlying all such studies is that the growth and development of the human organism is heavily influenced by nutrition status. This reflects the balance between food intake and energy expenditure. Nutritional status, in turn, is then used as a broad measure of living standards—with quality as well as quantity implications—that provides the rationale for the study of anthropometric data by nonmedical specialists.

However, whereas the work of Fogel and others focuses primarily upon the height of adult males as reflected in military records, Ward's study seeks to infer women's living standards from the weight of their newborn children. For adults, final height is strongly positively correlated with cumulative lifetime nutrition; for the newborn infant, birth weight is positively correlated with maternal nutrition prior to pregnancy and, more strongly, with maternal nutrition during pregnancy.

A substantial body of original research—birth weights and information about maternal characteristics such as age, occupation, length of hospital stay, and the number of preceding children borne to over 42,000 women from the middle of the 19th to the early 20th centuries—underlies Ward's study. Separate formulaic chapters are devoted to each city and to the hospitals and the patients they served. These are followed by a graphical presentation of the time-series data on average birth weight and a multi-

ple regression analysis of the data. The book concludes with a somewhat less successful chapter that tries to pull together the separate experiences in the five cities over time and to relate them to the broader theme of economic development. Unfortunately not every hospital collected the same data for their patients. As a result, cross-national comparisons of the influence of different characteristics on birth weight and maternal well-being is difficult. More important, though, the lack of references in the book to relative real income levels (even Angus Maddison's data in *Economic Growth in the West* [New York: Norton, 1964] would have been a useful addition), long-term economic growth, short-term economic fluctuations, and the role of income inequality in the determination of the level and distribution of nutrition prevents any real inferences about, or analysis of, the relationship between industrialization and women's living standards. Similarly, the very broad renderings of social class and occupation and the absence of information about maternal living arrangements make sociological interpretations of the status of women in these societies all but impossible from the data.

Of one thing, though, we are sure: The samples in Ward's study are biased. Only a small fraction of total births were delivered in hospitals at that time and their patients do not constitute representative cross sections of their urban populations. Consequently, one must be especially careful in generalizing from this study.

The hospitals in Ward's study catered primarily to the poor and to the unwed. It thus comes as a surprise to find that these mothers delivered remarkably sturdy babies, weighing an average of 3,200–3,500 grams (between seven pounds and seven pounds eleven ounces). Either these mothers lived well or the fetus is remarkably resilient and tolerant of a wide range of maternal conditions. One suspects the latter, though there is no evidence in the book to refute the former hypothesis.

Average birth weights, however, are just one statistic from underlying distributions that are pictured nowhere in this study. This oversight is unfortunate, for it would have been helpful in interpreting the data if we could have had a sense of how these historical distributions compare with and overlay those of more recent time periods, for which we have additional information on maternal nutrition and other indicators of the quality of life.

Ward's results contain one other surprise: Hospitals during that time, though seething dens of disease, improved maternal nutrition and well-being, since mothers with longer stays prior to delivery delivered heavier babies. The difference is small (indeed, one might argue, unimportant despite its *statistical* significance), two to four grams per day of hospital stay prior to delivery, but it does suggest that some interesting input-output calculations could be done using hospital diets to infer activity levels or diet outside of the hospital. Last, it would have been nice to know how these women and their children fared after birth, if only in the period prior to their discharge from the hospital. How many suc-

cumbed to disease? How many died? Such questions clearly impinge upon the notion of women's living standards.

Despite these shortcomings and lingering questions, Ward has grappled with a large body of new data to produce an original and important study that adds yet another piece to the standard-of-living-during-industrialization jigsaw puzzle but without necessarily fitting it into place.

*Contraception and Abortion in Nineteenth-Century America.* By Janet Farrell Brodie. Ithaca, N.Y.: Cornell University Press, 1994. Pp. xviii + 373. \$33.95.

Carl N. Degler  
*Stanford University*

Between 1800 and 1900, the birthrate of white women in the United States fell 49%, with 55% of that decline occurring between 1840 and 1880. Describing how that precipitous drop took place is the purpose of Brodie's book. Reduction in fertility among immigrant groups and African-Americans is not included in her study. Part of that exclusion results from the kinds of sources Brodie draws upon: some 60 now obscure medical and self-help books from the 19th century, medical journals, newspapers, and manuscript and archival records, almost all of which report information about white native-born Americans. Brodie's central purpose is to describe the methods used to control reproduction over the century and the ways those methods were disseminated.

Much of the book's information will be familiar to anyone who knows of earlier studies by other scholars on abortion and contraception, but, thanks to the author's close study of the 19th-century literature, Brodie's book provides richer detail than any previous book about the kinds and uses of the various methods of 19th-century contraception. Most students of the subject know that douching, the rhythm method, withdrawal, and condoms were used before the Civil War, but Brodie concludes that condoms were never as popular as the other methods because sheaths were associated with venereal disease and prostitution. Brodie rightly emphasizes that the spread of contraception depended not upon novel devices—if only because none was introduced until the diaphragm late in the century—but upon a socially determined effort to reduce fertility. New values, not new methods, supplied the key.

Brodie does an excellent job in reflecting on some of the ambivalences ordinary couples experienced as they sought to limit their fertility. Their frequent aim, Brodie notes, was simply to space children rather than to conclude reproduction, as modern couples often do. And how practical, she wisely asks, was douching in an unheated house without running water in the dead of winter? Brodie makes a strong case that small

towns, rural regions, and the working and artisan classes gained access to medical and contraception knowledge, usually from the mails (such mail-order information was more common then than now). Nothing gives a better idea of these subtle concerns than her first chapter, on the reproductive and sexual habits of a single middle-class couple. The chapter owes its detail to Brodie's imaginative "decoding" of the woman's diary. Two of the several valuable items it reveals are the strong sexual interest of these two Victorians and the anxiety that haphazard fertility control raised in the wife, and probably in many other women, given the earnest but always incomplete and often wrong information provided at the time. But that anxiety, too, is a measure of the drive to limit fertility.

Although Brodie clearly relates the push to reduce births to women's peculiar interest, she essentially agrees with the standard view that fertility control was a cooperative activity by both partners. She shows that the public proponents of controlling fertility were generally males, like the well-known Charles Knowlton and Robert Dale Owen and the more obscure Edward Bliss Foote, each of whom she treats fully. The opponents of birth control were also males, like Anthony Comstock and the great antiabortionist H. R. Storer. Women birth controllers, like Elizabeth Cady Stanton, were likely to be closeted with other women and to eschew print. In point of fact, the story of abortion in the United States is treated only in a single chapter on the opponents of reproductive control. (The subject is best treated in James Mohr, *Abortion in America: The Origin and Evolution of National Policy* [New York: Oxford University Press, 1978]).

By the nature of the book, a central question implicit in it cannot really be addressed, namely, When and why did the new values pressing for fertility control emerge? Nevertheless, this fine piece of research and analysis will surely be a part of the search for an answer.

*Inventing the Feeble Mind: A History of Mental Retardation in the United States.* By James W. Trent, Jr. Berkeley and Los Angeles: University of California Press, 1994. Pp. xii + 356. \$30.00.

Colin Samson  
*University of Essex*

In this scholarly and lucid book, James Trent provides us with a detailed social history of mental retardation in the United States. He has accomplished this task by what appears to be many years of hard work in libraries and archives. His narrative draws extensively upon institutional records, policy statements, photographs, visual imagery, letters, and other records of educationalists, psychologists, psychiatrists, parents, inmates, and patients. At the theoretical level, Trent draws upon both the Foucauldian concept of the gaze and social constructionism, although neither of these orientations are made explicit in the main body of the

text, which reads quite clearly as a social history. In my view, this is an advantage of Trent's work, as the narrative is not interrupted, like so many works in "deviance," by the literal application of these concepts to specific empirical observation. Rather, Trent is both more subtle and vivid, taking the reader through a chronology of the various constructions of mental retardation, relying heavily upon the publications and conference reports of the major ideologists from about the mid-19th century on to the present. Although Trent sometimes provides excessive renderings of minutiae, his interesting vignettes of the intellectual careers and personal lives of such key figures as Edward Seguin, perhaps the first "professional" in the field in the United States, the asylum superintendent Isaac Kerlin, and the eugenicist Henry Goddard add an immediacy to the book.

What emerges is an account of the complex and continuously changing nature of definitions, nomenclature, and classifications of mental retardation that is always related to the flux of professional and political interests. The first real authoritative voice on "idiots" in the United States came from pedagogy. It is in this world that Trent broadly locates Seguin's influential concept of physiological education, which held that the idiot required physical exercise, education of the senses, and attention to the will (p. 46) so that a "respectable mediocrity" could be achieved. With time, however, this relatively empathic emphasis gave way to administrative concerns and medicalization, which both posited more directly deterministic constructions of those designated as feeble-minded. Signaling a retreat from Seguin's relatively optimistic "mediocrity," by the late 19th century this determinism was reflected in more functional education, military drills, and custodial control (p. 83), as well as a general shift in the location of provision from schools to asylums.

A large portion of Trent's history deals with the persistent attempts from the late 19th century through to the Second World War to constitute mental defect as resulting from hereditary transmission. In perhaps the most impressive chapter of the book, on the menace of the feeble-minded, Trent weaves a web linking eugenics to American culture, immigration, race, and crime. These associations, put forward most notoriously in pedigree studies such as Goddard's *Kallikak Family*, were augmented by the contribution of the intelligence test from academic psychology, which was used to provide quantitative boundaries for classifying variations in mental pathology. The vogue of eugenics is associated, perhaps tacitly by Trent, with the cult of success in American society. Intellectual and material success are values so strongly believed in that those who have not achieved within these narrow confines have been assumed to be somehow *by nature* tainted, a proposition that has a contemporary ring about it, especially in some of the biomedical psychiatric discourse surrounding homelessness, as well as in sociobiology. For varying lengths of time, both total institutionalization and involuntary sterilizations were used in many states to check the growth in the numbers of feeble-minded. At a practical level, however, sterilizations were often

undertaken by professionals to help maintain social order in institutions, rather than out of eugenic zeal.

In the more recent period, Trent analyzes the legacy of the Kennedy community mental health legislation, which, he persuasively argues, had the effect of uniting several groups at the center of debates in mental health, mental retardation, and social policy—liberal critics of the abysmal conditions in institutions, labeling theorists who contested the veracity of mental pathology categories, reformers such as Wolf Wolfensberger, IQ determinists such as Arthur Jensen, and antiwelfare conservatives. The community policy outcome of this complex mixture of antiinstitutional voices has been not a deinstitutionalization but a move to other types of institutions in the community in the form of privatized intermediate care facilities. Other types of reinstitutionalization, such as imprisonment, have also appeared. In an aside, Trent reminds us that Bill Clinton authorized the execution of a feeble-minded Arkansas murderer in January 1992.

Perhaps, however, reinstitutionalization only underlines how ambivalent we must be about any talk of progress. What is needed, and in my view Trent is persuasive here, is not an analysis of mental retardation, but a skepticism toward mental acceleration, and “a recognition that in a society which defines and confines all meaning and worth in terms of production, profit and pervasive greed, intellectually disabled people will likely be exploited” (p. 277). This book is both sobering and stimulating. It would be a fitting accolade to it and its subject matter, if it helps to influence decision makers and academics and puts mental retardation back on the map.

*Deciding Who Lives: Fateful Choices in the Intensive-Care Nursery.* By Renee R. Anspach. Berkeley and Los Angeles: University of California Press, 1993. Pp. xii + 303. \$20.00.

Charles Bosk  
*University of Pennsylvania*

One of the intellectual pleasures of a tightly focused analytic ethnography is that it allows a close comparison between how actors describe their behavior and their actual behavior. Very good ethnography, then, allows us to see conflicts over meaning as they occur in action. One arena where such conflicts are particularly intense is the neonatal intensive care unit, where physicians, nurses, and parents frequently clash over what is “necessary treatment in a baby’s best interest.” Renee Anspach’s *Deciding Who Lives* is an exceptionally good ethnography that allows us to see how conflict arises, assent is produced, or dissent is defused in the intensive-care nursery.

Anspach’s work is based on 16 months of first-hand observation at two sites. The neonatal intensive-care unit studied in most depth is the

intensive-care nursery in an elite urban medical school, with all that that implies for being on the frontiers of neonatal medicine. The second site was the special care nursery of a public acute-care hospital of over 1,000 beds for the indigent.

From this rich experiential base, Anspach fashions a book composed of six chapters and two appendixes. The introductory chapter presents a brief synopsis of the research design and major organizing themes of the ethnographic narrative. The second chapter is a critical review of past theorizing on life and death decisions. In this area, as Anspach correctly points out, much is written but little is known. This chapter shows to full advantage Anspach's considerable theoretical skills. First, she lays out a powerful critique of a bioethical approach to life and death decisions: Such an approach focuses on "cases that are ethically complex but prognostically simple," emphasizes the principles rather than the process of decision making, and improperly assumes an individual decision maker (pp. 35-39). Then, Anspach argues for an "alternative perspective" that sees decisions as multimeaning transactions that emerge as both organizational and historical products. At a time when the bedside is so crowded with researchers, medical sociologists and ethnographers owe a clear debt to Anspach for so clearly and concisely spelling out what sociology brings to the discussion of medical decision making.

Chapter 3, "Predicting the Future," opens Anspach's discussion of her field materials. Neither the observations nor the analysis built upon them are original to Anspach. In interesting ways the very fact others have discussed the same phenomenon underscores Anspach's analytic strengths as she refines, extends, and sharpens our understanding of decision making on medical frontiers through a careful dissection of the sociology of clinical knowledge. This chapter focuses on how conflicts over an infant's prognosis between doctors and nurses are anchored in the routine work experiences of each. Physicians rely on "hard" scientific data and discount impressionistic data gathered at the bedside. Nurses, who, unlike physicians, have sustained contact with infants and derive much work satisfaction from that contact, use and trust interactional cues. This structuring of knowledge has two consequences. First, it encourages optimistic prognostic judgments among physicians and negative ones among nurses. More importantly, since nurses's judgments are built on data discounted by physicians, their judgments are not seriously weighed in the discussion of what should be done.

With so much at stake, with so many actors with different agendas, and with so much pressure from uncertainty, the widespread assent in intensive-care units is as remarkable as the occasional intractable conflict. In the next chapter, Anspach details the production of assent. This chapter is an important part of Anspach's dialogue with bioethics. It makes problematic the phenomenon of agreement, something that ethicists do not examine since their presence is only required during conflict. The analysis itself, which emphasizes the manner in which parents are excluded from decision making, is a discussion of how a cherished principle



of ethical treatment (informed consent) is subverted by workplace ideology, practices, and contingencies. Again, what is impressive is not the originality of the insight itself but how much Anspach adds to what is known.

The last (empirical) chapter to deal directly with the ethnographic material chronicles the diffusion of dissent. Three strategies dominate staff practices: anticipation and preemption, persuasion, and neutralization through psychologizing. To these, we might now add deflecting negative feeling through consultation with staff bioethicists. In truth, we do not know how these strategies are deployed, because ethnographic work on the ethical dilemmas of hospital medicine has yet to examine the everyday work of the in-hospital ethicist. Anspach's work provides an important starting point for hypothesizing about the impact of these new specialists who are no longer "Strangers at the Bedside."

A concluding chapter moves "Beyond the Nursery" and discusses how paradoxes of public policy create everyday dilemmas for those who provide neonatal intensive care. In addition there are two useful methodological appendixes: one discussing the field work and one the interviewing. This book deserves a wide audience both in and beyond sociology. In sociology, it is a wonderful example of how ethnography sharpens the understanding of public problems. A good read, this book has a place in introductory courses as a model of ethnographic work, in courses in the sociology of medicine as an exemplar of the sociology of medical knowledge, and in fieldwork courses as a how-to-do-it manual. Outside the discipline, this book is valuable for the explicit ways it compares and contrasts a sociological with a bioethical approach to medical practice.

*What Machines Can't Do: Politics and Technology in the Industrial Enterprise.* By Robert J. Thomas. Berkeley and Los Angeles: University of California Press, 1994. Pp. xvii + 314. \$45.00 (cloth); \$16.00 (paper).

Steven Peter Vallas  
*Georgia Institute of Technology*

In the 30 years since Robert Blauner's classic analysis of automation and factory life (*Alienation and Freedom* [Chicago: University of Chicago Press, 1964]), a welter of contending approaches toward work and technology has emerged, none of which seems satisfactory. Robert J. Thomas's ambitious *What Machines Can't Do* promises to change this situation. Drawing on qualitative case studies of workplace automation in four different manufacturing industries, Thomas has made a major contribution that should recast research in the field for some time to come.

In his opening chapter, Thomas develops a useful critique of the dominant perspectives on organizations and technology, clearing the ground for his proposed "power-process" theory. In his view, previous research

has largely missed the dynamic, processual elements involved in technological change. When these *have* been made the object of research (as in strategic choice theory), the focus has been on elite managerial decision makers, to the exclusion of other, nonelite employees. Thomas's power-process model aims to transcend these limitations, viewing technological change as the product of subtle processes of organizational contestation, in which both elite and nonelite actors bring their divergent worldviews to bear upon decisions involving new machines. This approach expects that not only corporate managers and elite design staff but also plant managers, local engineering personnel, and unionized workers will struggle to institutionalize their interests and ideals by using new technology as a lever with which to effect organizational change.

The substantive core of the book is developed in chapters 2–5, which present highly engaging accounts of technological change in four different branches of manufacturing: the aircraft, computer, aluminum, and automobile industries. The unit of analysis here is not the industry but discrete instances of technological change (e.g., flexible machining systems, computer numerical control, and robotic assembly systems). In each case, Thomas traces the decision-making process from its early stages, where organizational problems are first identified, through the selection of specific technologies to the point of actual implementation. Hence the book extends its inquiry backward in time and downward in the organization, thus breaking new ground.

Like labor process theorists, Thomas finds ample evidence of a schism between capital and labor. But unlike his neo-Marxist counterparts, Thomas situates class divisions alongside a host of other conflicts, many of which remain invisible when viewed through the labor process lens. The most important of these for Thomas is the organizational tension between product design engineers (who are typically employed in research and development divisions) and manufacturing or process engineers (usually employed in local plants). As he notes, the structure of U.S. industry accords product designers much greater power and status than their counterparts in manufacturing operations, confining the latter group within a rigid set of constraints. Deprived of autonomy and recognition, manufacturing engineers respond by advocating the adoption of new process technologies that would demonstrate their technical abilities and potentially recast their role within the firm. Such motives are scarcely apparent within written proposals or organizational surveys, of course, where a public discourse of efficiency prevails.

On some occasions, manufacturing engineers use formally sanctioned means to gain corporate approval, often by promising to bring allegedly inefficient or unruly departments under closer managerial control (as in the case of the aircraft industry). On other occasions, advocates of technological change resort to unauthorized means to accomplish their goals, even mobilizing quasi-insurgent movements to alter the structure of the firm (as in computer manufacturing). Perhaps the most interesting cases discussed in the book, however, are those in which plant managers and

manufacturing engineers come to form alliances with their workers, embracing collaborative models of technological change that dramatically alter the social relations of production (as occurred in aluminum casting and auto parts production and in the introduction of computer numerical control).

The question naturally arises as to why collaborative models emerge in some contexts rather than others. Drawing on Giddens's structuration theory, Thomas reasons that whether Fordist systems are reproduced or transformed depends in large part on the "propositions" (for Giddens, the "rules") that govern the process of technological change (pp. 247–49). Thomas suggests that open or inclusive processes are governed by "integrative propositions" that invite the participation of diverse occupational groups, thus favoring the transformation of Fordist organizations. Because closed or exclusive processes restrict participation to elite employees, however, they generate "zero-sum propositions" that instead tend to reproduce Fordist methods. This reasoning is suggestive but seems incomplete, for it fails to explain why the propositions governing change processes take one form or another and why they are sometimes subject to change in midstream.

The book contains other flaws. At times the power-process model seems to comprise not so much a theory as a set of orienting assumptions whose concrete expectations are not made entirely clear. Moreover, Thomas tends to focus one-sidedly on intraorganizational processes, at the expense of organizational fields and political influences that may well impinge on the processes at hand. These points notwithstanding, *What Machines Can't Do* has brought to light aspects of organizational life that have too often gone ignored and promises to open up new and important avenues for research on work, technology, and organizational change.

*Japanese Auto Transplants in the Heartland: Corporatism and Community.* By Robert Perrucci. New York: Aldine de Gruyter, 1994. Pp. xii + 186. \$37.95 (cloth); \$18.95 (paper).

Michael Schwartz

*State University of New York at Stony Brook*

This book invites comparison with Martin Kenney and Richard Florida's *Beyond Mass Production* (New York: Oxford, 1993), because they both focus on the creation of Japanese automobile manufacturing complexes in the American Midwest. But the two books are very different. Kenney and Florida are passionately interested in the internal structure and production methodology of the transplants and the possibility that the Japanese structure will become the heart of a transformed American automobile industry. Perrucci is interested in the exterior aspects of the process, focusing his attention on the ways in which the transplants related to the

various constituencies in the host communities to which they migrated. This focus, together with the book's modest size, makes Perrucci's book useful as a general introduction to transplants. It is accessible to nonspecialists and even advanced undergraduates, while still containing enough original material to engage researchers in the area.

Though it is organized around important empirical data, the most significant virtue of Perrucci's book is the ambitious integration of existing theories of corporate-community relationships. Chapter 2, which enunciates this argument, contains a useful review of corporatism and growth machine theory and presents a fresh idea, "embedded corporatism," as a substitute notion. As the book develops, however, it is clear that embedded corporatism is more of an extension of growth machine logic than anything else, and its relationship to Granovetter's concept of embeddedness is somewhat vexed. Nevertheless, Perrucci offers an arresting argument that growth machine politics can be an arena for channeling corporate interest into what he calls a new "social economy," a set of dynamics that encompasses nonelite interests in a progressive way.

The empirical chapters of the book present a series of discrete measurements focused on different aspects of the transplant process. Chapter 3 (coauthored with Fanying Kong), on the choice of location, makes two important contributions. First, it offers convincing evidence that the mid-western locations were largely a consequence of the availability of suppliers in near enough proximity to populate a just-in-time production system. (This finding has an important political consequence: It demonstrates the commitment of the Japanese assemblers to re-creating their system within the United States and validates their claims that they are committed to high levels of domestic content in their vehicles.) Second, it documents that the commitment of local government was far more significant than the dollar value of the incentive packages (or of wage concessions). This finding lays a foundation for Perrucci's argument that the Japanese were seeking to fully embed themselves in the local growth machine.

In chapter 4, Perrucci undertakes a careful content analysis of press coverage and concludes that the coverage was overwhelmingly positive and that dissenters were often ignored or relegated to back-page coverage. There are a number of lovely insights sprinkled through the chapter, including the almost ethnomethodological exploration of how apparently objective coverage of protest was transformed by the deft invocation of "experts" into support for the project. There is a missed opportunity in the coverage of opposition to Toyota in Lexington, Kentucky. Perrucci does not analyze this circumstance sufficiently well to offer a more general understanding of when such opposition will arise and when it will receive media coverage commensurate with its strength.

Chapter 5, on worker socialization, is nicely ambiguous. It provides empirical grist for the most cynical perspectives on the transplants by

revealing the antiunion and anti-Black biases of the recruitment process. But it also finds virtues in the work groups and other devices designed to boost worker input into the production process.

The final two chapters return to the concept of embedded corporatism and arrive at a deliciously paradoxical set of conclusions. On the one hand, Perrucci argues convincingly that the integration of Japanese transplants into midwestern communities has strengthened the dominance of business in these localities by making the local government a permanent partner in the campaign to introduce and sustain the Japanese production system. On the other hand, Perrucci finds considerable virtue in this new domination, because the process of constructing it "permits non-market measures of community development" to develop and become a part of the local economic trajectory of the locality (pp. 148-49). In the longer run, Perrucci argues, this allows for a new social economy derived from the "re-embedding of capitalism within the social institutions of the community" (p. 160). Though Perrucci explicitly denies it, this appealing vision is reminiscent of the early days of capitalism, when local entrepreneurs depended on the good will of local commerce and on a solidly organized working-class community (see, e.g., Herbert Gutman, *Work, Culture, and Society in Industrializing America* [New York: Knopf, 1976]).

Perrucci's vision is a compelling one, but it remains unproven. One point that gives me pause is his lack of documentation of structural embeddedness. Perrucci treats embeddedness as mainly a matter of positive attitudes spread across key segments in the community. If he could document that the transplant corporatism created material dependency on nonelite structures, I would find the argument considerably more persuasive. As it stands, once the transplants are established, they could lose their commitment to nonmarket measures of community development and disestablish their ties to social institutions in the community. If the only penalties for such disassociation are the disillusionment of the community, the transplants would be no worse off than the American automobile manufacturers were. While this might make them as vulnerable to new competition as the Big Three were to the Japanese invasion, there is no new competitor on the horizon.

Perrucci himself has doubts, and this dialectical ambivalence is one of the virtues of the book. As he says at the end of Chapter 5, "This may, in fact, be the beginnings of a new direction in economic life, one that is based on fairness, democracy and community. Or it may simply be business as usual, with large corporations using a new way to gain the cooperation of government and labor while they pursue their basic goal of profit" (p. 123).

*Defending Community: The Struggle for Alternative Development in Cedar-Riverside.* By Randy Stoecker. Philadelphia: Temple University Press, 1994. Pp. xv + 304. \$44.95 (cloth); \$18.95 (paper).

Alford A. Young, Jr.  
*University of Chicago*

In *Defending Community: The Struggle for Alternative Development in Cedar-Riverside* Randy Stoecker depicts the tribulations of a community mobilization effort designed to ward off an urban renewal initiative. By the late 1960s, Cedar-Riverside had become a counterculture community that was regarded by many of its inhabitants as "the Haight-Ashbury of the midwest." The remarkable feature of this effort was that Cedar-Riverside managed to maintain much of its counterculture flavor well into the 1980s (by then the author took up residence in the community while pursuing his doctoral studies at the University of Minnesota). This study is a result of Stoecker's interviews with community residents and activists, archival research on community institutions and organizations, and assessments of the impact of municipal, state, and federal involvement in community development for Cedar-Riverside from the mid-1960s until the mid-1980s.

This book is most generally a contribution to the sociology of the urban community. More specifically, Stoecker addresses issues and perspectives such as the status of culture in community-level social organization, the attempt to manage conflict resolution in community organizing, and sociological perspectives on community development. One of the emphases of this work is the analysis of the Cedar-Riverside community as a site of resistance. Stoecker argues for a modification of decentralized segmentary reticulate theory (DSR), which says that successful movements emerge from the mobilization of many disparate groups, each maintaining different members, that foster similar or overlapping agendas. He labels his modification to DSR "federated frontstage structure," which refers to the same individuals creating or joining different organizations that are targeted as specific objectives during a protest engagement. This additional concept allows Stoecker to distinguish between social organizational approaches that may be more effective for community-level mobilization and those more appropriate for larger scale social movement activity.

Stoecker also addresses the debate in new social movement theory in his exegesis on how collective (in this case, community) identity facilitates the interests and goals of community inhabitants during particular points in time. He synthesizes this focus with his assessment of community development activity, where he explores how local, state, and federal entities conceive of and execute plans for community change. In essence, the community residents desired to preserve a quasi-countercultural residential space for themselves, while the federal- and state-funded developers wanted to garner greater economic return than what existed previ-

ously in that space. The residents won most of the battles with the developers (acceding only a portion of the community to the creation of high-rise housing). This result was achieved because Cedar-Riverside's "residents were young, highly educated, and thoroughly radicalized. They (also) had an abundance of specialized, home-grown talent in urban planning, architecture, and financial management, as well as time to devote to the community" (p. 247).

An additional emphasis in this work concerns the effort of community residents to move from the position of activists working in the interests of preserving their desired type of community to managers of that community once they successfully fought off much of the urban renewal initiative. In this circumstance, two community groups took opposing sides in the effort to negotiate the conflicting interests of community control of neighborhood institutions, resources, and local culture against that of creating mechanisms for economic growth that would benefit individual residents and families. In describing how these two ends promoted tension in the efforts of the community to manage its own affairs, Stoecker says, "A structural economic contradiction became an inter-organizational conflict" (p. 189). By considering the effects of such role revisions, Stoecker provides critical insight into the relationship of social activism to the end-state of social change.

Rather than serving as an in-depth case study of a particular phenomenon (such as urban social movement theory or the sociological perspective on community development), *Defending Community* reads more like an inquiry into a unique historical phenomenon that, through Stoecker's critical analysis, sheds light on a range of issues pertinent to some of the most sociologically relevant matters in urban studies. No single theoretical approach is completely validated or enhanced by Stoecker's mode of inquiry into this case. Instead, this work provides an entrée into the possible directions or avenues that a study more intentionally focused on one or a few theoretical approaches may take. Consequently, the reader is left demanding greater emphasis on some specific questions, theoretical implications, and policy discussions that are introduced in the book. However, Stoecker's choice of putting the uniqueness of this case ahead of any particular theoretical issue or substantive problem in urban community research leaves the reader with a rich commentary that links together all of the major questions, and many of the little ones, that beset any situation involving neighborhood change and the range of social actors who have at least some stake in that effort for change.

*Contemporary Urban Sociology*. By William G. Flanagan. Cambridge: Cambridge University Press, 1993. Pp. 185. \$49.95 (cloth); \$14.95 (paper).

Harvey M. Choldin

*University of Illinois at Urbana-Champaign*

The purpose of William Flanagan's short book is to guide the reader through the theoretical controversies that have ripped urban sociology since 1972, when Manuel Castells first blasted human ecology as its dominant theory. The author's presentation is extraordinarily well organized, which makes his arguments easy to follow, despite their complexities. He presents a series of debates over alternative theories, showing how each debate shifts over time and what its turning points were. This organization proves to be a most effective way of structuring a book about theories. The book's tone is consistently positive; Flanagan emphasizes the ways in which the various theories are useful.

Altogether, *Contemporary Urban Sociology* is an outline of the field, a guidebook to the issues and literature. The volume includes a thorough, up-to-date bibliography in which 178 of the 260 references were published since 1980. Like the book in general, the bibliography incorporates many references from the field of social geography.

I am doubtful about one aspect of the volume, the brief sections called "Managing Urban Lives" or "Managing Urban Spaces," with which the author concludes each chapter. In each case Flanagan takes a specific issue, gentrification for example, and shows how the chapter's theoretical issues apply to it. There is a disjunction between the bodies of the chapters and these appendages; in each case, the example incorporates only a couple of ideas from a preceding exposition of greater complexity.

Chapter 1 deals with community life in the urban context. Flanagan discusses two types of studies, those concerned with social networks, which tend to emphasize spatially dispersed relationships, and those studying neighborhoods, which tend to emphasize social activities within specific areas. He relies upon neighborhood studies published in the 1980s, concluding that there is vitality in this sort of research, but then he neglects it for the rest of the book, which is mostly written at the macrolevel.

Chapter 2 is about human ecology. While Flanagan takes human ecology seriously, I was disappointed by his neglect of ecological research on residential segregation by race. In my estimation this line of research represents one of (urban) sociology's proudest accomplishments. It includes a series of studies conducted over a period of more than 30 years, employing an effective and consistent methodology, and reveals fundamentally important social facts and trends. To me these strong studies show that the contribution of human ecology lies more in its research products than in its theory.

I particularly liked chapter 3, "Urban Political Economy and Its Critics." The chapter begins by explaining why urban sociology encountered



a crisis in the decade of the 1970s. Flanagan explains that the old paradigm was inadequate to answer many of the central social issues of the times: "The drama and urgency of the major social upheavals taking place in the urban arena remained out of focus for urban ecology. . . . A new paradigm was needed to satisfy the call for an urban science relevant to the hour" (p. 74). That new approach was political economy. Flanagan explains political economy as a theoretical framework that emphasizes the world system with its capital structures and their effects. For urban sociology this means that any given locality is secondary to the larger system. The body of chapter 3 elaborates upon and presents critiques of various versions of this framework. Flanagan is most appreciative of Gottdeiner and Feagin's point of view, which stresses the study of individual cities within the larger context in order to discover how particular communities are created—a combination of the global with the local. In general, chapter 3 explains well the transformation of urban sociology.

Chapter 4, entitled "The Third World and the World System," surveys the topic of Third World urbanization, showing that the new framework revitalizes this topic, which used to be a sort of appendage to urban sociology. Flanagan shows that all the old topics—underdevelopment, overurbanization, primate cities, and the like—all gain new meaning and importance within the new perspective.

Finally, in chapter 5 Flanagan argues that urban sociology must address a new agenda "to discover what it is that makes each city unique in its response to global forces, and to understand how it is that some cities are able to resist general regional trends while others typify them" (p. 137). Flanagan says that we must reemphasize human agency. At this point the author brings in Anthony Giddens's theory of structuration, saying that it offers a future direction for urban sociology. Criticizing urban sociologists for concentrating on the macrolevel, Giddens enjoins them to look at the interplay between social action and social structure. Referring specifically to four of Giddens's writings that discuss cities and the effects of urbanism (published between 1981 and 1989), Flanagan encourages urban sociologists to study Giddens and to use his ideas.

Surprisingly, this upbeat book concludes with paragraphs that express doubt about the viability of urban sociology as a field of inquiry. While he is willing to make his choices about the most productive directions for the field to take, Flanagan does not go so far as to say that the viability of urban sociology is a certainty.

This is a rich and fascinating book. I enjoyed reading it, and it taught me a great deal. It is worth rereading and will be a valuable teaching aid for graduate students and probably for undergraduates as well. Flanagan has done a great service in writing this fine book.

*Homeless in Paradise: A Map of the Terrain.* By Rob Rosenthal. Philadelphia: Temple University Press, 1994. Pp. ix + 265. \$44.95 (cloth); \$18.95 (paper).

Daniel M. Cress

*University of Colorado at Boulder*

Rob Rosenthal's *Homeless in Paradise* is a refreshing departure from most treatments of the homeless issue. He takes aim not only at accounts that emphasize disaffiliation and personal pathology but at those whose structural emphasis implies that the homeless are passive victims of socio-economic forces beyond their control. He does not reject the presence of these factors, but rather, he reinterprets their significance in light of the experiences of the homeless who resided in Santa Barbara during the mid-1980s.

Rosenthal depicts a highly heterogeneous population that is in part a function of his use of a broader definition of homelessness. He argues convincingly that the critical factor for homeless people is whether they have a home base they can depend upon, and he therefore defines as homeless those who lack a legal residence of their own. Thus, his study includes those who are forced to double up with others to avoid landing on the streets as well as the literal homeless. The emphasis on the "invisible" homeless encompasses groups such as women and Latino families that have been underrepresented in other studies of the homeless and is an aspect that distinguishes this work from other research on the issue. Thus Rosenthal's portrayal is not so much a portrait of the homeless as it is a family album that includes an array of homeless types that differ from one another in their survival routines but who share the same desire for a place of their own.

One main thrust of the work takes issue with the disaffiliation thesis, which asserts that the homeless are detached from larger social institutions and each other. This theme was central to studies of skid row a generation ago, and it appears implicitly in modern treatments that emphasize personal disability and pathology. But rather than reject this premise outright, Rosenthal argues that disaffiliation varies by the type of homeless and operates as a consequence of homelessness rather than a cause. Thus, certain types of homeless are more integrated than others, having stronger connections to family, work, social service agencies, and each other.

Rosenthal is much more sympathetic to structural arguments, particularly those that emphasize the decline in affordable housing, the lack of jobs that pay a living wage, and cutbacks in social service programs for the poor, as explanations for the rise in homelessness. Yet he rightly criticizes the implication in other works that document the same processes that the homeless are passive victims of these forces. Having spent almost five years with the homeless in Santa Barbara, Rosenthal documents a story of survivors who struggle against the insanity of the social welfare

system, who spend long hours working or looking for work, or who are simply trying to get food, shelter, and perhaps a shower. To be sure, there is plenty of despair and resignation among his informants, but they are hardly inert casualties. Rather, they grapple with these broader forces in their daily routines in an attempt to secure some level of stability in their lives.

The critiques of the disaffiliation thesis and the passive victim implication are demonstrated most compellingly in a chapter on homeless political action. Mobilization occurred initially among the most detached homeless type, the street people, in response to police raids on their encampments. They formed the Homeless People's Association and engaged in disruptive protest with the city of Santa Barbara over policing and sleeping policies and voting rights. They also engaged in a public relations campaign to counteract negative stereotypes, emphasizing the long-term residency of many of their members, their desire for work, and their status as veterans. Homeless women formed a second organization, the Single Parent Alliance, which focused on issues of concern to homeless families. Thus, Rosenthal finds a high level of organization among the supposedly disorganized as well as collective patterns of resistance to their plight.

There are some drawbacks to this work. The chapter on becoming homeless breaks no new ground and has a banal flavor. In addition, the treatment of some of his homeless types, particularly street kids and Latino families, could be more developed. Relatedly, the work begs for some tables to succinctly characterize how his homeless types vary on the dimensions with which he is concerned. Finally and most perplexing, there is nothing in the book about how the experience of homelessness might be different in "paradise" settings. While the title implies an exceptional location for the presence of homelessness, Rosenthal plays up only the commonalities between his setting and others. Yet Santa Barbara is not the Cass Corridor of Detroit or North Philadelphia. Thus, he plays for generality but sacrifices a potential contribution of a unique setting. These criticisms notwithstanding, *Homeless in Paradise* is a fine piece of work that makes thoughtful and original contributions to our understanding of the homeless phenomenon.

*Constructing White-Collar Crime: Rationalities, Communication, Power.* By Joachim Savelsberg, with contributions by Peter Bruhl. Philadelphia: University of Pennsylvania Press, 1994. Pp. xiv + 179. \$29.95.

James William Coleman  
*California Polytechnic State University*

This book is a case study of the passage of a law most Americans have probably never heard of: the Second Law against Economic Crime, which was enacted in the Federal Republic of Germany in 1986. But

even though this law is an obscure one to American readers and Savelsberg doesn't present any startling new findings or theoretical breakthroughs, this book is a solid piece of work full of interesting insights. Those concerned with white-collar crime or the sociology of law will, I think, find a perusal of this volume worth the effort.

Savelsberg's methodological approach to this case study was to construct a sequential series of cognitive maps showing the way actors involved in the problem of economic crime conceptualized it and showing their assumptions about the causal relationships underlying the problem. His raw data came from documents and interviews with key individuals, which were then analyzed by two independent coders. Analyzing these cognitive maps in time sequence reveals the patterns of change in the way the issue of economic crime was constructed as it moved from a generalized social concern to actual legislation.

Savelsberg divides this process of law formation into three phases. The first begins with "increasingly massive claims making by mass media, criminal justice practitioners, politicians, and academics" (p. 8). Savelsberg analyzes the way the German press constructed the general problem of white-collar crime and goes on to explore the viewpoints of the legal profession and industry associations. The second phase began with the appointment of an expert commission to look into the problem, and Savelsberg analyzes its cognitive assumptions and the recommendations it made. Finally, he explores the perspectives of the various political actors and industry lobbyists involved in the enactment of the final bill, which ended up to be a watered-down version of the legislation recommended by the expert commission. The most important difference is the elimination of key provisions to criminalize antitrust offenses such as bid rigging.

Perhaps the greatest strength of this book is the way Savelsberg carefully draws out the theoretical implications of his case study for the sociology of law and our understanding of white-collar crime legislation. I particularly appreciated the balanced appraisal he brings to each of the theoretical perspectives he explores and the way he avoids the unsubstantiated generalizations in which more heated advocates of a particular theoretical position often indulge. His carefully qualified theoretical conclusions are not easily summarized in a short review, but one central theme does emerge. Savelsberg argues in several places, using several different lines of evidence, for the overall superiority of what he calls the action/conflict-group orientation to functionalist approaches. However, Savelsberg uses the rubric of functionalism more broadly than many others, for he includes both conservatives who see the law as a reflection of the functional necessities of society as a whole and Marxists who see it as a reflection of the needs of the capitalist class. While not denying the importance of such forces, Savelsberg argues that the general needs of society or the economy are not, in and of themselves, sufficient to account for the enactment of new legislation and that the activation of specific interest groups and the personal involvement of individual actors are also necessary before change can occur. Moreover, the final result of

this process is not always in the long-term interests of society or the capitalist class as a whole but may reflect the particularistic interests of the groups and individuals directly involved in the legislative struggle.

Savelsberg concludes his book with a comparison of the evolution of white-collar crime legislation in Germany and the United States that might be of particular interest to American readers. Unfortunately, the German example provides little hope for those of us who have been frustrated by the half-hearted and ineffective response to the problem of white-collar crime in the United States. Savelsberg argues that "critical public sentiment and debates on the behavior of the powerful is not just older but more persistent in the United States than in Germany. It further seems that government action in seasons of public mistrust has also been somewhat more forceful in the United States than in Germany" (p. 159). For these differences, he offers several explanations that revolve around the centralized nature of German society in which the legal profession, politicians, and the media are closely tied to centralized power structures and group conflicts are carried out largely through "neocorporate" interest organizations rather than grassroots political activity.

*Poor Discipline: Parole and the Social Control of the Underclass, 1890–1990.* By Jonathan Simon. Chicago: University of Chicago Press, 1993. Pp. x + 286. \$15.95 (paper).

John R. Sutton  
*University of California, Santa Barbara*

Since its inception at the end of the 19th century, the institution of parole has muddled along in the shadowy margins of the prison system. Its stated function was to ease the movement of convicts from prison to community, but revisionist scholars have suggested that its main functions were legitimation and surveillance: Parole gave a soft therapeutic gloss to the modern penal regime while expanding its capacity for control. Rehabilitation went out of public fashion in the get-tough 1970s, and for awhile it looked as if the precipitous growth of U.S. prisons would render parole superfluous. Jonathan Simon informs us that parole has survived quite nicely by radically adapting its structure and practices to a new and more punitive climate.

In fact, argues Simon, parole has re-formed itself several times, whenever economic and social changes raised new challenges for the prison system. Parole was not originally meant to be therapeutic, and Simon shows that it became attractive to legislators only when prison-labor schemes broke down and rising populations of idle inmates threatened to make prisons unmanageable. Parole forged an external link between the prison and the labor market: The promise of a job was the primary criterion for parole release, and parole officers spent much of their time finding jobs for ex-convicts. In the 1950s a combination of deindustrial-

ization and rising drug use made this disciplinary model untenable, and parole shifted to a clinical model that gradually decentered the role of labor. Official rules continued to insist that parolees work, but in practice parole officers were satisfied to place them in the growing network of drug treatment and welfare programs.

The political and legal changes of the 1970s brought about a further change to what Simon calls "penal postmodernism." Determinate sentencing laws and court rulings sharply curtailed the discretion of parole agents, and the hardening of the urban underclass severed the link between the prison and the labor market. The task of parole is no longer to rehabilitate nor to reintegrate but to manage a perpetually jobless and crime-prone population according to policies that are increasingly technocratic and self-referential. As Simon demonstrates, parole revocations accounted for an uncontrollably large proportion of prison admissions in the 1980s; with the decline of faith in the modern penal project, penalty has become, in his terms, "power without narrative" (chap. 8).

Simon builds his analysis on an exceptionally broad range of data sources: historical studies, government reports (with special attention to California), national and local criminal justice statistics, and the author's participant observations in an Oakland parole office. He uses these data shrewdly to advance a complex argument that operates at several levels of analysis. His observations from the field are impressive enough on their own, but they gain considerable heft from his attention to the political and economic environments within which parole agents do their work. *Poor Discipline* is a convincing piece of empirical sociology.

I am less convinced, however, by the theoretical inferences Simon draws from his findings—specifically, his claim that "postmodern" penalty is a qualitatively new development. Imprisonment has always been a mixed project, successful precisely because it could accommodate a range of punitive and therapeutic motives within a single (albeit loosely articulated) discursive frame. And parole, as Simon suggests (p. 35), has always been an organizational "garbage can," an administrative solution in search of a problem. Decades of research have shown that criminal justice officials are more concerned with practical managerial concerns—clearing cases, avoiding controversy, staying within a budget—than with ideological zealotry. I find myself wishing that Simon had paid more attention to the recent literature on organizations, where many of the characteristic features of criminal justice agencies he identifies—the symbolic nature of formal rules, the decoupling of structure and practical action, the role of institutional environments, and the tendency to emphasize procedural rather than substantive measures of effectiveness—have been extensively theorized.

Given the remarkable adaptive capacity of the criminal justice system, it might be premature to declare the death of penal modernism; indeed, criminal justice officials are increasingly suggesting that we have gone too far in a punitive direction. But Simon might be right to suggest that this momentum is too powerful to reverse. The combination of the decline

of the economy, the withering of the welfare state, and the social quarantining of urban minority communities has created a helpless and irresistible source of prison fodder, and there is no sign that these collateral problems are going to be addressed soon. Further, to cite a development that came too late for Simon's book, the California prison guards' union has emerged as a powerful force behind the Draconian "three strikes and you're out" initiative. The prison system is becoming, like James Joyce's image of Ireland, a sow that eats her own farrow.

# DURKHEIM ON RELIGION

W. S. F. Pickering, editor

New Low Price  
for Classroom Adoption



This volume brings together the most comprehensive selection yet available of Émile Durkheim's writings on religion. Besides helping to establish the discipline of sociology in France, Durkheim is widely recognized as one of the founding figures in the modern study of religion. The volume contains a new translation of important sections of *The Elementary Forms of Religious Life* (1912) and a selection of Durkheim's early reviews, articles and extracts from books, presented in order of

original publication. Also included are comprehensive bibliographies and early reviews of Durkheim's work on religion by such writers as van Gennep, Goldenwieser, Stanner, and, for the first time in English, by the French sociologist Gaston Richard.

Code: 01 02 06

Paper: \$19.95

ISBN: 1-55540-981-4

Scholars Press • P.O. Box 6996 • Alpharetta, GA 30239-6996  
1-800-437-6692



## JUST PUBLISHED



NEW EDITION

### SOCIETY AND TECHNOLOGICAL CHANGE

Third Edition

RUDI VOLTI, *Pitzer College*

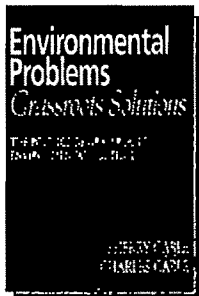
Presents an overview and analysis of the causes and consequences of technological change

■ Considers social, economic, political, and cultural aspects ■ Facilitates understanding of

subject matter by covering topics of particular interest to students  
■ Provides an international perspective

1995 / 320 pp. / Paper / ISBN 0-312-09642-9

NEW



### ENVIRONMENTAL PROBLEMS/GRASS-ROOTS SOLUTIONS

The Politics of Grassroots Environmental Conflict

SHERRY CABLE, *University of Tennessee-Knoxville*

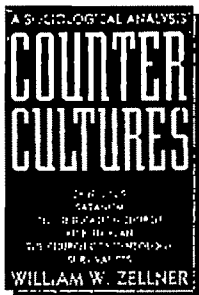
CHARLES CABLE

A narrative text that examines how sociologists identify, analyze, and address political and economic issues of environmental conflict

■ Presents environmental problems as social, not technical, problems  
■ Fills a gap among textbooks for environmental sociology courses  
■ Contains an environmental science primer for non-hard science students and instructors

1995 / 158 pp. / Paper / ISBN 0-312-08142-1

NEW



### COUNTERCULTURES

A Sociological Analysis

WILLIAM W. ZELLNER, *East Central University*

Examines six countercultures – Skinheads, Ku Klux Klan, Survivalists, Satanic Cults, the Church of Scientology, and the Unification Church ■ Author's participant observation with groups connects methodology to actual research

■ An innovative text written for an academic audience

1995 / 213 pp. / Paper / ISBN 0-312-08084-0

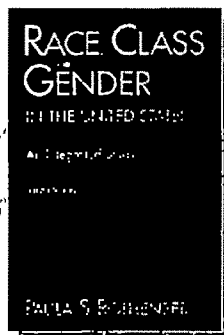


St. Martin's Press • College Division  
175 Fifth Avenue, New York, NY 10010

1-800-446-8923

Fax: (212) 780-0115

# FROM ST. MARTIN'S PRESS



*NEW EDITION*

## **RACE, CLASS, AND GENDER IN THE UNITED STATES**

**An Integrated Study**

*Third Edition*

**PAULA S. ROTHENBERG**, *The William Paterson College of New Jersey*

Examines issues of race, gender, and sexuality within the context of class society in the United States ■ Employs a cross-disciplinary approach with readings

from scholarly essays, non-fiction, fiction, and poetry ■ Readings are interesting and relevant, enabling instructors to initiate lively discussions among students ■ Allows students to better understand their ordinary experiences within a conceptual framework

1995 / 523 pp. / Paper  
ISBN 0-312-09652-6

*NEW EDITION*

## **THE ELEMENTS OF SOCIAL SCIENTIFIC THINKING**

*Sixth Edition*

**KENNETH HOOVER** and **TODD DONOVAN**, *Western Washington University*

A basic introduction to the vocabulary, concepts, and objectives of social scientific thought ■ Readable and engaging style provides students with a common-sense, rather than technical, approach to grasping major concepts ■ Accessible to students with no prior statistical background ■ Cross-disciplinary examples will interest students from a wide variety of areas in the social sciences

1995 / 224 pp. / Paper  
ISBN 0-312-09643-7



**St. Martin's Press • College Division**

175 Fifth Avenue, New York, NY 10010

1-800-446-8923

Fax: (212) 780-0115

S T . M A R T I N ' S P R E S S

# **TEACHING OVERSEAS:**

**The Department of Social and Behavioral Sciences at the American University of Beirut (AUB) in Beirut, Lebanon invites applications for three faculty positions: two in the field of Psychology (Social/Personality and Experimental/Research Design), and one in the field of Sociology, all starting October 1, 1995. Rank will normally be assistant professor; usual contract is for three years. The Department offers both bachelor's and master's degrees. Candidates will be expected to engage in research and supervise student research and thesis writing. The language of instruction is English.**

**Applicants should hold the Ph.D. in Psychology or Sociology, with postdoctoral teaching and/or research experience required, and are expected to teach both undergraduate and graduate courses in the Department.**

**Interested candidates should send a letter of application (with a copy to the Director of Personnel, c/o AUB NY Office) and a c.v., and should arrange for three letters of reference to be sent, before February 15, 1995, to: Dean, Faculty of Arts and Sciences, American University of Beirut, c/o AUB New York Office, 850 Third Avenue, 18th Floor, New York, NY 10022-6297, USA. Incomplete and/or late applications will not be considered.**

**AUB is an Affirmative Action/Equal Opportunity Employer.**

**U.S. passports are presently invalid for travel to, in or through Lebanon, and for residence in Lebanon, by order of the Department of State, and therefore applications from individuals who would travel to or reside in Lebanon on a U.S. passport cannot at this time be considered.**

Mouton de Gruyter

# Linguistic Human Rights

## Overcoming Linguistic Discrimination

*Edited by Tove Skutnabb-Kangas and Robert Phillipson,  
in collaboration with Mart Rannut*

This book describes what linguistic human rights are, who enjoys them and who is deprived of them and why, and suggests which linguistic rights should be regarded as basic human rights.

Discrimination against language minorities is widespread, despite much national and international law which prohibits it. Generally it is only speakers of official languages (speakers of dominant majority languages) who enjoy all linguistic human rights. This volume documents and analyzes rights in many countries worldwide.

It is relevant for sociolinguists, language planners, educationalists, teachers, human rights lawyers, activists, and governmental and non-governmental organizations. The book addresses issues such as whether speakers of minority languages can use their mother tongue in school, in their dealings with administrative or legal authorities, and in participation in democratic processes. It provides a wealth of documentation, supported by a theoretically informed conceptual framework for the analysis of linguistic human rights worldwide and their promotion in the future.

A substantial appendix contains extracts from selected U.N. and regional documents covering linguistic human rights, proposals for such, and resolutions on language rights.

1994 • vi + 478 pages • \$145.00 • ISBN 3-11-014370-4

*Contributions to the Sociology of Language 67*

Price subject to change

*For North America:*

**Walter de Gruyter, Inc.**

200 Saw Mill River Road Hawthorne, NY 10532

Tel: (914) 747-0110 • Fax: (914) 747-1326

Berlin • New York

# MONOGRAPHS

of the Society for Research in Child Development

## **The Development of Sex Typing in Middle Childhood**

*Lisa A. Serbin, Kimberly K. Powlishta, and Judith Gulko*

How do children associate sex roles with occupations, activities, and traits? As they develop sex-typed models, are they expressing personal preferences, learned social stereotypes, or both? This volume presents the results of interviews with more than 500 children and shows the impact of cognitive maturation, family environment, and socialization on the development of sex typing. "Provides a new perspective for viewing the development of gender differences in schooling and work and offers strategies for reducing gender disparities." — *Marlaine E. Lockheed*.

\$14.00 0-226-74870-7

(58:2, serial no. 232, 1993)

## **Coping with Marital Transitions**

*E. Mavis Hetherington and W. Glenn Clingempeel*

"A landmark contribution to the research literature on divorce" — *Eleanor Maccoby*. Almost half of the children born in the last decade will experience the divorce of their parents, and many of these will also go through the changes associated with their custodial parent's remarriage. How do parents and their children adjust through the stages of marriage, divorce, single parenthood, and remarriage? How does adolescent development vary among intact families, single parent, and remarried families? This longitudinal study examines family relationships (including parent-child, marital, and sibling) and adolescent development during these important transitions.


\$18.50 0-226-33168-7 (57:2-3, serial no. 227, 1992)

Published for the Society for Research in Child Development by

**The University of Chicago Press**

Journals Division, P.O. Box 37005, Chicago, IL 60637

Fax 312/753-0811



**THIS  
PUBLICATION  
AVAILABLE  
FROM UMI**

This publication is available from UMI in one or more of the following formats.

- In Microform—from our collection of over 18,000 periodicals and 7,000 newspapers
- In Paper—by the article or full issues through UMI Article Clearinghouse
- Electronically, on CD-ROM, online, and/or magnetic tape—a broad range of ProQuest databases available including abstract-and-index, ASCII full-text, and innovative full-image format

Call toll-free 800-521-0600, ext. 2888, for more information, or fill out the coupon below:

Name \_\_\_\_\_

Title \_\_\_\_\_

Company/Institution \_\_\_\_\_

Address \_\_\_\_\_

City/State/Zip \_\_\_\_\_

Phone ( ) \_\_\_\_\_

I'm interested in the following title(s) \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

UMI  
A Bell & Howell Company  
Box 78  
300 North Zeeb Road  
Ann Arbor, MI 48106  
800-521-0600 toll-free  
313-761-1203 fax

**UMI**

Ingwer Borg and  
Peter Mohler (Editors)

## Trends and Perspectives in Empirical Social Research

1994. viii + 504 pages.

Paper \$53.95 3-11-014312-7

Cloth \$114.95 3-11-014311-9

**Trends and Perspectives in Empirical Social Research** is a comprehensive and authoritative volume on the state of the art in a wide range of approaches, methods, and techniques, contributed by an international group of senior scholars from all branches of empirical social research. The volume includes chapters on social indicators, nationwide general social surveys, multinational surveys, cognitive and communicative aspects of survey measurement, official/census data, computer-assisted interviews, sampling, statistics and measurement, scaling, facet theory, causal modeling, attitude measurement, multilevel models, phenomenology, hermeneutic approaches, and content analysis.

*Prices subject to change.*



For North America:

**Walter de Gruyter, Inc.**  
200 Saw Mill River Road  
Hawthorne, NY 10632



## **Repertoires and Cycles of Collective Action**

MARK TRAUGOTT, EDITOR

*Contributors.* Charles D. Brockett, Craig Calhoun, Doug McAdam, Marc Steinberg, Sidney Tarrow, Charles Tilly, Mark Traugott, James W. White  
208 pages, 7 tables, 13 figures, paper \$13.95,  
library cloth edition \$28.95

## **A Rhetoric of Bourgeois Revolution**

The Abbé Sieyès and *What is the Third Estate?*

WILLIAM H. SEWELL, JR.

This volume studies the powerful rhetoric of *What is the Third Estate?* and the brilliant but enigmatic thought of its author.

224 pages, paper \$11.95, library cloth edition \$36.95  
*Bicentennial Reflections on the French Revolution*

## **Duke University Press**

Box 90660 Durham, North Carolina 27708-0660

# **Public Opinion Quarterly**

Stanley Presser, Editor

### **Examining methods and theories of**

survey validity  
interviewing and interviewers  
mode of administration

questionnaire construction  
sampling strategy  
analytic approaches

### **Covering issues that include**

political tolerance  
gender consciousness  
AIDS awareness  
social identity

voter behavior  
environmental consciousness  
cross-national comparisons  
crime reporting  
media threats

If you are a political scientist, sociologist, psychologist, journalist, polling director, statistician, market researcher, or communications specialist, you will find that *POQ* is an indispensable resource.

Published for the American Association for Public Opinion Research  
by The University of Chicago Press

Regular one-year subscription rates: \$23.00 Individuals, \$20.00 Students (with copy of valid ID), \$50.00 Institutions. Outside USA, add \$4.00 for postage. MasterCard and Visa payments accepted. Members of the American Association for Public Opinion Research receive *POQ* as part of their membership benefits; write for more information about AAPOR.

To order your subscription, send check, purchase order, or complete credit card information to The University of Chicago Press, Journals Division, Dept. SF55A, P.O. Box 37005, Chicago, IL 60637.

# GAME THEORY FOR POLITICAL SCIENTISTS

JAMES D. MORROW

Game theory is the mathematical analysis of strategic interaction. In the fifty years since the appearance of von Neumann and Morgenstern's classic *Theory of Games and Economic Behavior*, game theory has been widely applied to problems in economics. James Morrow's book is the first to provide a standard text adapting contemporary game theory to political analysis. It uses a minimum of mathematics to teach the essentials of game theory and contains problems and their solutions suitable for advanced undergraduate and graduate students.

This rigorous but accessible introduction to game theory will be of use not only to political scientists but also to sociologists and others in the social sciences.

Cloth. \$35.00 ISBN 0-691-03430-3

# THE SOVEREIGN STATE AND ITS COMPETITORS

HENDRIK SPRUYT

The present international system, composed for the most part of sovereign, territorial states, is often viewed as the inevitable outcome of historical development. Hendrik Spruyt argues that there was nothing inevitable about the rise of the state system. Examining the competing institutions that arose during the decline of feudalism, Spruyt disposes of the familiar claim that the superior size and war-making ability of the sovereign nation-state made it the natural successor to the feudal system.

Spruyt's interdisciplinary approach not only has important implications for change in the state system in our time, but also presents a novel analysis of the general dynamics of institutional change.

*Princeton Studies in International History and Politics*

Cloth: \$39.50 ISBN 0-691-03356-0

PRINCETON UNIVERSITY PRESS

AVAILABLE AT FINE BOOKSTORES OR DIRECTLY FROM THE PUBLISHER: 609-883-1759



# CHICAGO

## Teen Mothers



Ruth Horowitz

### Teen Mothers—Citizens or Dependents?

Ruth Horowitz

"[This book] demonstrates how large-scale bureaucratic rationalities are mistranslated into daily policies through face-to-face interactions and suggests how emotionally intelligent work could restore to welfare recipients their sense of dignity and help them find a place of active citizenship in society."

—Dmitri N. Shalin, University of Nevada, Las Vegas

"In the first study of inner-city teen welfare mothers to probe the civic-moral dimension at the heart of the issue, not only does Horowitz analyze welfare as a problem of social control, but her analysis demonstrates the best of feminist theory—redressing a telling silence so directly and pragmatically as to compel a radical revision to received wisdom."

—Mark D. Jacobs, George Mason University

Cloth \$32.95 264 pages

### Now in Paper Getting a Job

*A Study of Contacts and Careers  
Second Edition*

*With a new Preface and an Afterword  
on the recent work in the field*

Mark Granovetter

This classic study of how 282 men in Newton, Massachusetts, found their jobs not only proves "it's not what you know but who you know," but also demonstrates how social activity influences labor markets.

"Who would imagine that a book with such a prosaic title could pose such provocative questions about social structure and even social policy?"

—Edward O. Laumann,

*American Journal of Sociology*

Paper \$13.95 264 pages

### Mama Might Be Better Off Dead

*The Failure of Health Care  
in Urban America*

Laurie Kaye Abraham

"A book of unexpected power. What begins as a matter-of-fact account of a black family's struggle to get medical care builds into an emotional indictment of our incomprehensible, illogical health care bureaucracy."

—Walt Bogdanich,  
*New York Newsday*

Paper \$10.95 297 pages

## THE UNIVERSITY OF CHICAGO PRESS

5801 South Ellis Avenue  
Chicago, Illinois 60637



0002-9602(199501)100:4;1-4

# AJS

*AV 12*

American  
Journal  
of Sociology  
Volume 100 Number 5  
March 1995

*Married Women's Employment in Rapidly Industrializing  
Societies—Brinton, Lee, and Parish*

*Marital Disruption and Mortality—Lillard and Waite*

*Organizational Context of Criminal Sentencing—Dixon*

*Structural Sources of Socioeconomic Segregation in Metropolitan  
Brazil—Telles*

*Role-Based Ecology of Technological Change—  
Podolny and Stuart*

## **SYMPOSIUM ON APPLIED REGRESSION**

*Methods for Comparing Regression Coefficients between  
Models—Clogg, Petkova, and Haritou*

*Comment: Impact of Random Predictors—Allison*

*More on Comparing Regression Coefficients—Clogg and  
Petkova with Cheng*

**The University of Chicago Press**

MARTA TIENDA, Editor

JIM DAVIS, FRED KNISS, SUSAN E. MAYER, JEFFREY MORENOFF, EMILIO PARRADO, and ROBERT J. SAMPSON, Associate Editors

MARY C. BRINTON, Associate Editor, Book Review

GRACE KAO, Assistant Book Review Editor

CHAN-UNG PARK, STEPHEN B. PLANK, MICHAEL J. REAY, DAVID A. SMILDE, and ALFORD A. YOUNG, JR., Book Review Editorial Board

SUSAN ALLAN, Managing Editor

LAUREN FORD-TAGGART and CLAIR JAMES, Editorial Assistants

**Consulting Editors** PAUL D. ALLISON • EDWIN AMENTA • PETER BEARMAN • ROBERT J. BURSİK, JR. • JOHN CASTERLINE • NANCY DAVIS • NANCY DENTON • CLAUDE FISCHER • ROBERTO FRANZOSI • ADAM GAMORAN • JOHN R. HALL • MICHAEL T. HANNAN • MICHAEL HECHTER • DOUGLAS HECKATHORN • BARBARA HEYNS • CHARLES HIRSCHMAN • K. JILL KIECOLT • TIMUR KURAN • GARY LAFREE • DAVID LATTIN • LINDA D. MOLM • DANIEL MONTI • JEYLAN T. MORTIMER • DANIEL OLSON • DUDLEY POSTON • CECILIA RIDGEWAY • CATHERINE ROSS • RUBÈN G. RUMBAUT • JOACHIM J. SAVELSBERG • DARREN E. SHERKAT • SHELLY SMITH • C. MATTHEW SNIPP • CHARLOTTE STEEH • GILLIAN STEVENS • ROBIN STRYKER • EDWARD TELLES • YU XIE

**International Consulting Editors** JOHN FOX (Canada) • MATTHIJS KALMIJN (The Netherlands) • NOAH LEWIN-EPSTEIN (Israel) • STEVEN L. RYTINA (Canada) • YUVAL YONAY (Israel)

**Editorial Board** ANDREW D. ABBOTT • GARY S. BECKER • CHARLES E. BIDWELL • DONALD J. BOGUE • RONALD BURT • TERRY N. CLARK • JAMES S. COLEMAN • JOHN L. COMAROFF • ROGER V. GOULD • WENDY GRISWOLD • EVELYN KITAGAWA • EDWARD O. LAUMANN • DONALD N. LEVINE • ALBERTO PALLONI • WILLIAM L. PARISH • MOISHE POSTONE • MARTIN RIESEBRODT • EDWARD SHILS • GEORGE P. STEINMETZ • ROSS M. STOLZENBERG • FRED L. STRODTBECK • GERALD D. SUTTLES • LINDA J. WAITE • WILLIAM JULIUS WILSON • KAZUO YAMAGUCHI

THE AMERICAN JOURNAL OF SOCIOLOGY (ISSN 0002-9602) is published bimonthly in July, September, November, January, March, and May by The University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, Illinois 60637. Subscription rates, U.S.A.: institutions, 1 year \$96.00, individuals, 1 year \$41.00. Student subscription rate, USA: 1 year \$28.00 (photocopy of valid student ID must accompany subscription). ASA and BSA members, 1 year \$36.00. All other countries add \$9.00 for each year's subscription to cover postage. Canadian subscribers add 7% GST. Subscription agent for Japan: Kinokuniya Company, Ltd. Single copy rates: institutions \$16.00, individuals \$6.85. Volumes 1-90 available from Schmidt Periodicals GmbH, Dettendorf, D-83075 Bad Feilnbach, Germany. All volumes are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, and Kraus Reprint and Periodicals, Route 100, Millwood, New York 10546. Subscriptions are payable in advance. Subscriptions will be entered to start with the first issue to be published after receipt of order. **Postmaster:** Send address changes to *American Journal of Sociology*, The University of Chicago Press, Journals Division, P.O. Box 37005, Chicago, Illinois 60637.

ADVERTISING space in THE AMERICAN JOURNAL OF SOCIOLOGY is available, as is rental of its subscriber list. For information and rates, please contact the advertising sales staff, The University of Chicago Press, Journals Division, 5720 South Woodlawn Avenue, Chicago, Illinois 60637. Advertising and list rental is limited to material of scholarly interest to our subscribers.

*Copying beyond Fair Use.* The code on the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made beyond those permitted by Section 107 or 108 of the U.S. Copyright Law provided that copies are made only for personal or internal use or for the personal or internal use of specific clients and provided that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc., Operations Center, 27 Congress St., Salem, Massachusetts 01970. To request permission for other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale, kindly write to Permissions Department, The University of Chicago Press, 5801 S. Ellis Ave., Chicago, Illinois 60637. If no code appears on the first page of an article, permission to reprint may be obtained only from the author.

Second-class postage paid at Chicago, Illinois, and at additional mailing offices

© 1995 by The University of Chicago. All rights reserved.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. (C)

8/10/96  
P-27A/A+CL

Sociology

AJS

*American Journal of Sociology*

Volume 100 No. 5 March 1995

## CONTENTS

- 1099 Married Women's Employment in Rapidly Industrializing Societies: Examples from East Asia  
MARY C. BRINTON, YEAN-JU LEE, AND WILLIAM L. PARISH

- 1131 'Til Death Do Us Part: Marital Disruption and Mortality  
LEE A. LILLARD AND LINDA J. WAITE

- 1157 The Organizational Context of Criminal Sentencing  
JO DIXON

- 1199 Structural Sources of Socioeconomic Segregation in Brazilian Metropolitan Areas  
EDWARD E. TELLES

- 1224 A Role-Based Ecology of Technological Change  
JOEL M. PODOLNY AND TOBY E. STUART

### **Symposium on Applied Regression**

- 1261 Statistical Methods for Comparing Regression Coefficients between Models  
CLIFFORD C. CLOGG, EVA PETKOVA, AND ADAMANTIOS HARITOU

- 1294 The Impact of Random Predictors on Comparisons of Coefficients between Models: Comment on Clogg, Petkova, and Haritou  
PAUL D. ALLISON

- 1305 Reply to Allison: More on Comparing Regression Coefficients  
CLIFFORD C. CLOGG AND EVA PETKOVA, WITH THE ASSISTANCE OF TZUWEI CHENG

### **Commentary and Debate**

- 1313 The Study of Social Taste through First Names: Comment on Lieberman and Bell  
PHILIPPE BESNARD

- 1317 Reply to Philippe Besnard  
STANLEY LIEBERSON
- 1325 Interclass Segregation, Poverty, and Poverty Concentration: Comment on Massey and Eggers  
CRAIG ST. JOHN
- 1333 Beyond the Technical Details: Reply to St. John  
DOUGLAS S. MASSEY AND MITCHELL L. EGGERS

### **Book Reviews**

- 1336 *Paradoxes of Gender* by Judith Lorber  
ROSANNA HERTZ
- 1338 *A Very Social Time: Crafting Community in Antebellum New England* by Karen V. Hansen  
JEANNE BOYDSTON
- 1340 *Encounters with Aging: Mythologies of Menopause in Japan and North America* by Margaret Lock  
L. L. CORNELL
- 1341 *Masculinity and the British Organization Man since 1945* by Michael Roper  
MARY BLAIR-LOY
- 1343 *Cinderella Goes to Market: Citizenship, Gender, and Women's Movements in East Central Europe* by Barbara Einhorn  
BARBARA HEYNS
- 1346 *The New American Workplace: Transforming Work Systems in the United States* by Eileen Appelbaum and Rosemary Batt  
VICKI SMITH
- 1347 *The Change of a Lifetime: Employment Patterns among Japan's Managerial Elite* by John C. Beck and Martha N. Beck  
HIROSHI ISHIDA
- 1349 *Union of Parts: Labor Politics in Postwar Germany* by Kathleen A. Thelen  
COLIN CROUCH
- 1351 *The Scar of Race* by Paul M. Sniderman and Thomas Piazza  
DAVID O. SEARS
- 1354 *Sharing the Journey: Support Groups and America's New Quest for Community* by Robert Wuthnow  
DAVID M. HUMMON

- 1356 *Goodness Personified: The Emergence of Gifted Children* by Leslie Margolin  
SPENCER E. CAHILL
- 1358 *Leaving Home before Marriage: Ethnicity, Familism, and Generational Relationships* by Frances K. Goldscheider and Calvin Goldscheider  
WILLIAM G. AXINN
- 1360 *Out of the Garden: Toys and Children's Culture in the Age of TV Marketing* by Stephen Kline  
GARY ALAN FINE
- 1362 *Careers and Creativity: Social Forces in the Arts* by Harrison C. White  
DIANA CRANE
- 1364 *All Things to All People: The Catholic Church Confronts the AIDS Crisis* by Mark R. Kowalewski  
JIM WOLF
- 1365 *The Social Control of Religious Zeal: A Study of Organizational Contradictions* by Jon Miller  
ADAM B. SELIGMAN
- 1367 *The Problem of Order: What Unites and Divides Society* by Dennis H. Wrong  
SATOSHI KANAZAWA
- 1369 *The Frankfurt School: Its History, Theories, and Political Significance* by Rolf Wiggershaus  
DOUGLAS KELLNER
- 1371 *The Birth of Fascist Ideology: From Cultural Rebellion to Political Revolution* by Zeev Sternhell, with Mario Sznajder and Maia Asheri  
STEPHEN TURNER
- 1374 *A Social History of Truth: Civility and Science in Seventeenth-Century England* by Steven Shapin  
KEVIN L. BROWN
- 1375 *Durkheim and Women* by Jennifer M. Lehmann  
EDWARD A. TIRYAKIAN
- 1378 **Book Notes**

## IN THIS ISSUE

MARY C. BRINTON is associate professor of sociology at the University of Chicago. Her current research projects focus on comparative gender stratification in East Asia and job search and recruitment processes in the United States, Japan, and South Korea.

YEAN-JU LEE is a researcher at the East-West Center. Recently, she has been investigating how rapid economic growth affects gender and generational relationships in East Asia.

WILLIAM L. PARISH is professor of sociology at the University of Chicago. His current research includes studies of aging in Taiwan and family and work organization in rural China.

LEE A. LILLARD is senior economist and director of the RAND Center for the Study of Aging and adjunct professor of economics at the University of Southern California. His research interests include socioeconomic determinants of mortality, economic consequences of poor health in older populations, health economics, labor economics, econometrics and models of marriage, cohabitation, marriage dissolution, and marital and non-marital fertility.

LINDA J. WAITE is professor of sociology of the University of Chicago, a research associate with the Population Research Center of NORC, and a research associate at RAND. She is president-elect of the Population Association of America. Her research has focused on the social demography of the family, including cohabitation, marriage, divorce, childbearing, and women's employment. With Frances K. Goldscheider, she coauthored *New Families, No Families: The Transformation of the American Home*, which received the 1993 Duncan Award for the best monograph in demography from the population section of the American Sociological Association.

JO DIXON is assistant professor in the Department of Sociology and the Institute for Law and Society at New York University. Her research interests include gun ownership, the role of organizational contexts in structuring plea bargaining and sentencing, spousal violence, and gender stratification in the legal profession. Her current research focuses on the nexus among sex, violence, and the state.

EDWARD E. TELLES is associate professor of sociology at the University of California, Los Angeles. His recent research has focused on the effect of economic development on race relations in Brazil. With Vilma Ortiz he is currently studying ethnic identity and mobility among Mexican Americans in Los Angeles and San Antonio.

JOEL M. PODOLNY is assistant professor of organizational behavior in the Graduate School of Business at Stanford University. His current research interests include the significance of status dynamics for market competition and technological change and the implications of social networks for mobility and satisfaction within organizations.

TOBY E. STUART is a doctoral student in organizational behavior in the Graduate School of Business at Stanford University. His areas of study include formal organizations and networks. He is currently researching the role of interorganizational networks in shaping change and corporate technology strategy.

CLIFFORD C. CLOGG is professor of sociology and statistics at Pennsylvania State University. His research interests include the social demography of the labor force and statistical models for the analysis of categorical data.

EVA PETKOVA is assistant professor of biostatistics in psychiatry at Columbia University. Her research includes developing methodology for statistical modeling in behavioral sciences. She is currently involved in a neurological study of suicide.

ADAMANTIOS HARITOU is professor in the Department of Applied Information at the University of Macedonia. He spent the 1993–94 academic year at Pennsylvania State University. His research interests include generalized linear models and statistical computing.

PAUL D. ALLISON is professor of sociology at the University of Pennsylvania. His recent papers include "Logit Models for Sets of Ranked Items" (*Sociological Methods* 1994; with Nicholas Christakis) and "Using Panel Data to Estimate the Effects of Events" (*Sociological Methods and Research*, November 1994). In addition to his methodological work, he is interested in theories of altruistic behavior and models of cultural evolution.

PHILIPPE BESNARD is Director of Research at the Centre National de la Recherche Scientifique, professor at the Institut d'Études Politiques de Paris, and a member of the Groupe d'Étude des Méthodes de l'Analyse Sociologique (C.N.R.S. and University of Paris—Sorbonne). His major research interests include Durkheim and the Durkheimians, suicide, social rhythms, and the social diffusion and transformation of tastes.

STANLEY LIEBERSON is currently writing a monograph on the social underpinnings of taste and the factors leading to changes. He is primarily using a wide variety of data on first names—current and past—to address these questions. The study is supported by a grant from the National Science Foundation. He is affiliated with Harvard University, where he is the Abbott Lawrence Lowell Professor of Sociology.

CRAIG ST. JOHN is associate professor of sociology at the University of Oklahoma. His current research projects examine the exposure of minority children to concentrated urban poverty and the effect of community context on unwanted pregnancy and material health-related behaviors during pregnancy.

DOUGLAS S. MASSEY is Dorothy Swaine Thomas Professor of Sociology at the University of Pennsylvania. He is currently carrying out research on the causes and consequences of international migration in Mexico in collaboration with Jorge Durand of the University of Guadalajara.



MITCHELL L. EGGERS is a research demographer at the United Nations Economic Commission for Europe's Population Activities Unit in Geneva, Switzerland. His major research interests include racial stratification and segregation in large urban areas of the United States and causes and consequences of population aging in Europe and North America.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

## Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures will be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract, on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for \$15.00, payable to The University of Chicago Press in U.S.

currency or its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is:

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 1/93)

# **Married Women's Employment in Rapidly Industrializing Societies: Examples from East Asia<sup>1</sup>**

Mary C. Brinton  
*University of Chicago*

Yean-Ju Lee  
*East-West Center*

William L. Parish  
*University of Chicago*

A variety of explanations have addressed the phenomenon of secular change in married women's employment in rapidly industrializing countries. These include theoretical frameworks that emphasize female labor supply, the conditions of labor demand, patriarchal values, the international division of labor, and the effects of export-led industrialization. This article examines two societies (South Korea and Taiwan) that showed considerable similarity in female labor supply conditions, female labor force participation, and cultural values 20 years ago but have since diverged in dramatic and puzzling ways. Using aggregate and microlevel data, this article shows that the emergent differences in married women's employment are best explained by the intersection of labor supply (similar in the two cases) and demand (markedly different). The article highlights the impact of government policy and foreign loan investment in shaping the nature of labor demand during rapid export-led industrialization in both countries.

## **INTRODUCTION**

The long-term secular increase in married women's economic participation in the 20th century across a range of countries has been a central

<sup>1</sup> Support from the Spencer Foundation, the National Institute of Child Health and Human Development (grant R01 HD23322), and the Chiang Ching Kuo Foundation is gratefully acknowledged. We also thank the Korean Women's Development Institute for allowing us to use the Korean data. Charles Chi-hsiang Chang provided research assistance for the paper. We received helpful comments from participants of the Labor Workshop and the Demography Workshop at the University of Chicago and from several *AJS* reviewers. Address all correspondence to Mary C. Brinton, Department of Sociology, 1126 E. 59th Street, University of Chicago, Chicago, Illinois 60637.

focus of much recent research in sociology and economics (Goldin 1990; Mincer 1962; Shimada and Higuchi 1985; Smith and Ward 1985). This article critically evaluates several explanations of the sources of change in married women's economic participation. We take advantage of a semicontrolled "natural experiment" provided by two societies—South Korea and Taiwan—that began the push for industrialization roughly two decades ago with many similar initial conditions. Despite their similarities, these societies subsequently diverged in their incorporation of women into the economy. By the 1980s, married women in the two societies showed surprising differences in their age patterns of employment, types of employment, returns to education, and rates of pay compared to those of men.<sup>2</sup>

The combination of similar initial conditions with differing outcomes in a pair of societies provides a unique opportunity to examine the roles played by the interaction of labor supply, labor demand, patriarchal values, the international division of labor, and export-led economic industrialization in structuring married women's changing labor force participation. This article focuses on developing an explanation that takes into account the interaction of social and economic forces at both the domestic and international levels. This is especially important for an understanding of the economic role of women in societies in Asia, Latin America, and Eastern Europe that have recently industrialized in the context of an increasingly internationalized world economy. We begin by discussing existing modes of explanation, then evaluate them using macro- and microlevel data on South Korea and Taiwan.

#### POSSIBLE EXPLANATORY MODELS

The *labor supply* conditions that lead to increased female labor force participation have been well theorized: as age at marriage increases and fertility falls with industrialization, decreasing family obligations free more of women's time for labor market activities. Increases in human capital (both education and work experience) also generally make women more attractive employees and raise their potential wage rate, which pulls more women into the labor market (Goldin 1990). This type of explanation leads to the hypothesis that differences in the quantity and quality (i.e., wages and working conditions) of women's employment in Taiwan and South Korea hinge primarily on characteristics of women's labor supply: education, prior work experience, and fertility levels.

<sup>2</sup> For examples of the earlier literature on the patterns in each society, see Kim (1990), Cho and Koo (1983), KWDI (1986), Chang (1982), Liu (1984), Chiang and Ku (1985), ROC (1991b), Galenson (1992), and Kang (1993).

*Labor demand* explanations suggest, however, that whether labor supply conditions such as increasing education and decreasing family obligations are translated into labor force participation depends on the nature of local labor markets (see, e.g., Chant 1991). In developing countries with low labor demand, women with more education may disdain the few lower paying, less prestigious jobs that are offered, leading to either an inverse or U-shaped relationship between women's education and work (Standing 1981; Smock 1981). Or, if demand for white-collar jobs remains modest, employers may reserve these jobs for men in order that they may earn a family wage. For example, in the United States and Britain before World War II and in Ireland until recently, strong "marriage bars" excluded educated married women from many white-collar jobs (Oppenheimer 1970; Goldthorpe 1987; Goldin 1990; Mincer 1962; Smith and Ward 1984; Pyle 1990).

A promising way of integrating supply-based and demand-based explanations is to analyze the nature of labor queues and women's place within them (Reskin and Roos 1990; Thurow 1969); that is, we may theorize that only when labor demand exceeds the supply of males in the queue will old prejudicial barriers to married women's employment fall. This type of explanation would lead to the hypothesis that differences between South Korea and Taiwan depend primarily on the level of labor demand in each society.

Explanations of married women's labor force participation that give primacy to *patriarchal values* suggest that, regardless of women's qualifications or the nature of labor demand, a myriad of cultural practices channel women into behaviors that either discourage labor force participation or encourage participation in only those jobs with the lowest income and prestige rewards (see, e.g., Papanek 1990). East Asian cultures such as Taiwan and South Korea are generally considered to be prime candidates for these types of explanations because of their strong extended-family, patrilineal descent traditions (see, e.g., Greenhalgh 1985; Salaff 1981; Jones 1984). Patriarchal values can affect behavior not only on the supply side but on the demand side as well, with policymakers, employers, and fellow employees making investment, hiring, training, and mentoring decisions that favor males over females (see, e.g., Pyle 1990). The hypothesis that emerges from this type of explanation is that differences between Taiwan and South Korea may be traced mainly to stronger patriarchal values in one of the two.

*New international division of labor* explanations typically argue that women at the periphery of the world economic system get the worst jobs in that system (see reviews in Pyle [1990], Tinker [1990], Wolf [1990], and Beneria and Feldman [1992]). One formulation of this idea places the blame for a large informal sector of makeshift, self-employed, service

jobs on multinational corporations that make inappropriately capital-intensive investments in developing countries (e.g., Frobels, Heinrichs, and Kreye 1980; Evans and Timberlake 1980; Kentor 1981; Timberlake 1985; London 1987). Another formulation emphasizes how menial, labor-intensive jobs are increasingly shifted from the core to the periphery of the world system, leading to a "fool's gold" explosion of jobs reserved only for young, single women who are dismissed as soon as they get married and begin having children (e.g., Nash and Fernandez-Kelly 1983; Sassen 1988). Either framework leads to the hypothesis that societies that are highly involved in the world economy produce inadequate demand for good jobs and this leads to employers' preference for young, unmarried women who will work for a few years at low wages and produce competitively priced goods for the world economy.

Finally, *export-led growth* explanations are popular in economic and World Bank circles. These offer the more optimistic view that suggests that unless local governments distort domestic factor markets (labor, capital, etc.), export production for the world market promotes labor-intensive jobs that soak up so much underemployed labor that employers must eventually raise wages and hire married as well as unmarried women (see, e.g., Fields 1985; Galenson 1992; Krueger 1983; Lim 1990). Only countries that put barriers on imports and exports, who hem in labor with inappropriate legislation, or who mistakenly subsidize capital fail to enjoy these benefits. This type of explanation suggests the hypothesis that economies that have higher levels of export involvement and little government distortion of domestic factor markets will generate more and higher-quality opportunities for married female workers.

#### METHODOLOGICAL APPROACH

We assess the applicability of these competing frameworks with both macro- and microlevel data on Taiwan and Korea. Our macrolevel analysis attempts to explain society-level differences in urban married women's employment. The microlevel analysis attempts to explain specifically *which* married women get better-quality employment. We therefore combine a discussion of the intersocietal differences in women's employment that need to be explained with a discussion of how the theoretical frameworks outlined above either facilitate or hinder understanding.

We use macrolevel data from commonly available surveys in the two societies and microlevel data from specialized surveys on married women's work. The macrolevel data, mainly from annual government surveys, provide the context for understanding patterns in the microlevel data. The microlevel data are from two nationally representative surveys: the Taiwan Women and Family Survey carried out in 1989 in

## Married Women's Employment

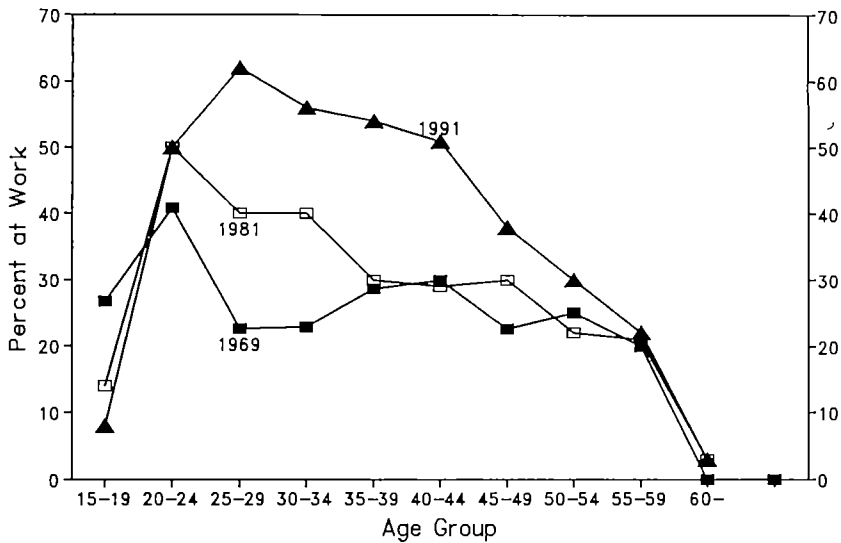


FIG. 1.—Women at work, Taipei City (ROC 1970, 1980, 1989)

Taiwan and the Survey of Women's Employment carried out in 1985 in South Korea. The Taiwan sample includes ever-married and never-married women 25–59 years old, while the Korean data are for ever-married women 15–59 years old. From each sample we select currently married women ages 25–49 who live in urban areas.<sup>3</sup> The emphasis on young and middle-aged married women in urban areas allows us to provide a fine-grained account of the opportunities and difficulties women face in the evolving labor market.

### WHAT NEEDS TO BE EXPLAINED

Taiwan and South Korea have shared many conditions in common, including extremely rapid export-led growth that has placed them among the leading “miracle” economies of East Asia. Given many similarities in their economic growth patterns and world system positions, the existence of three major differences in women's labor force participation are puzzling. First, during the early years of export-led industrialization, female labor force participation was very similar in South Korea and Taiwan (figs. 1 and 2, bottom line). In 1970, when one-third to one-half

<sup>3</sup> To correct for a slight underrepresentation of large cities, the Taiwan data were weighted to accurately reflect the distribution of the urban population across different city sizes.



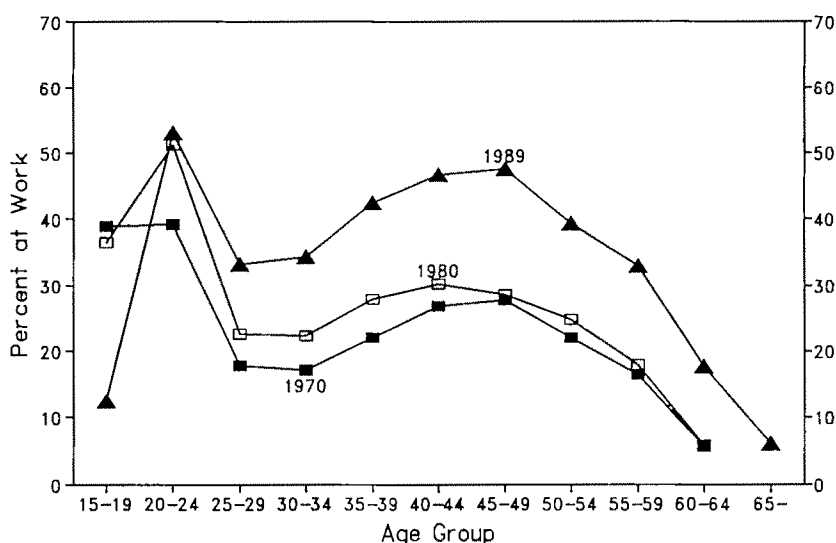


FIG. 2.—Women at work, urban Korea (ROK 1969, 1981, 1991a)

of the labor force remained in agriculture, women in both societies exhibited a pattern of delaying marriage while they worked. This was typically followed by a postmarital exit from the labor force. Despite these initial similarities in women's labor market behavior, the Korean and Taiwan patterns have since diverged (figs. 1 and 2, top lines). Older Korean women have increasingly entered the labor force once their children start school, a behavior that parallels and even exceeds similar patterns in Taiwan. But in great contrast to Taiwan, many Korean women in the prime childrearing ages of 25–34 years remain outside the labor force (see also Cho and Koo 1983).

Second, there are differences not only in the *quantity* of work for young married women but also in the *quality* of work available to all married women. Our microlevel data reinforce the macrolevel data in showing that the age-specific rates of labor force participation for married women are much lower in South Korea than in Taiwan (table 1). Korean married women ages 25–29 participate in the labor force at only one-half the rate of their Taiwan counterparts (27.1% compared to 54.9%), and women's labor force participation rates remain lower in Korea than in Taiwan for older age groups as well. Note that this is mainly a result of the fact that Korean married working women are much less likely to be in formal employment (i.e., working as employees) than are women in Taiwan. Why do married women's rates of formal employment at any age in Taiwan so exceed their rates in South Korea? This is a puzzle.

## Married Women's Employment

TABLE 1

### EMPLOYMENT STATUS DISTRIBUTION AMONG URBAN MARRIED WOMEN BY AGE

AGE GROUP (in years)	TAIWAN				SOUTH KOREA			
	Self-		Family	Total	Self-		Family	Total
	Employee	Employer	Enterprise Worker		Employee	Employer*	Enterprise Worker	
25-29 ... ..	36.7	13.0	5.3	54.9	8.1	12.5	6.5	27.1
30-39 ..... ..	29.6	18.3	10.1	58.0	13.8	17.0	8.1	38.9
40-49 .....	30.0	19.8	7.8	57.6	19.8	15.8	11.6	47.2
All ages .....	30.6	18.1	8.8	57.5	13.3	15.4	8.6	37.3

SOURCES.—Taiwan Women and Family Survey (1989), Survey of Women's Employment (1985).

\* Among Korean women who reported themselves as employees, those who work at their own house (i.e., domestic out-workers) are reclassified as self-employed.

Third, although we save detailed statistics until later, it is striking that *the consequences of education* differ in the two societies. In Taiwan, higher levels of female education lead uniformly to a higher probability of employment. In South Korea, just the opposite occurs. Among Korean married women, more education leads to a lower probability of employment, thereby violating standard models that predict that higher human capital leads to the promise of higher income, which in turn attracts women into the labor force.

These, then, are the three major differences that emerged for married women in the two societies by the mid-1980s: lower overall labor force participation in South Korea (most strikingly for young married women); lower rates of formal employment (employees working for wages or salaries) in South Korea; and the failure of higher human capital to lead to higher probabilities of employment in South Korea. Macrolevel data provide initial clues as to the applicability of the competing explanatory models outlined at the beginning of the paper.

### MACROLEVEL ANALYSIS

#### Labor Supply Conditions

Taiwan and South Korea have shared many similar labor supply conditions (table 2). In 1970, mean age at first marriage was slightly lower in Taiwan, but the gap between the two societies narrowed and then disappeared entirely by the end of the 1980s. Likewise, the crude birth rate plummeted in both societies. A parallel rise in marriage age provided more opportunities for women to accumulate work experience before marriage in both countries, and a shortened childbearing period made it

TABLE 2  
LABOR SUPPLY CONDITIONS

	1970	1975	1980	1985	1990
Family obligations:					
Age at first marriage, female:					
Taiwan .....	22.2	22.3	23.4	24.4	25.6
Korea .....	23.3	23.6	24.1	24.8	25.5
Crude birthrate.					
Taiwan .....	27.2	23.0	23.4	18.0	16.6
Korea .....	29.5	24.6	23.4	16.4	15.6
Education.					
12-17-year-olds in secondary education (%):					
Taiwan:					
Male .....	NA	69.5	71.9	77.9	83.6
Female .....	NA	61.7	70.0	78.6	87.4
Korea:					
Male .....	51.4	65.9	86.1	91.0	94.3
Female .....	33.1	48.5	76.1	87.0	90.5
18-21-year-olds in higher education (%).					
Taiwan:					
Male .....	NA	11.2	11.9	14.2	18.3
Female .....	NA	8.7	10.2	13.5	19.6
Korea:					
Male .....	13.9	13.9	25.2	51.0	50.8
Female .....	4.7	5.6	8.7	23.2	24.4

SOURCES.—Taiwan: ROC (1991b, 1991c); Korea: KSA (1991b), ROK (1990, 1991).

NOTE.—NA signifies that statistics were not available for this year.

possible for women to return to work sooner. Rising educational attainments also provided the conditions for an increase in women's labor force participation in both societies. Levels of female education have been slightly different; a higher proportion of women in Taiwan than in South Korea enrolled in secondary education, but Korean university enrollments overtook those in Taiwan. (Men's education patterns in Taiwan and Korea diverged, an important point to which we return later.) Examining female life-cycle patterns alone, the common phenomena of increasing age at marriage, declining child care obligations, and increasing education would lead labor supply explanations to predict not *diverging* but *converging* labor patterns for married women in the two societies. Thus, attention to labor supply conditions is clearly insufficient to explain the societal differences in women's labor force outcomes.

### Patriarchal Values

Similarities in patriarchal values in the two societies also render incomplete a value-based explanation. Both societies share a common Confu-

cian cultural tradition that emphasizes patrilineal and patrilocal family patterns. Until recently, it was common for women in Taiwan and South Korea to delay marriage while contributing some of their income to their family of birth and then to exit the labor market at least temporarily to concentrate on childrearing and household chores (Greenhalgh 1985; Jones 1984). Though there of course may be some distinctions between Taiwan and South Korea in sex-role values, the two societies have been quite similar (see, e.g., Arnold and Kuo 1984; Lee 1985). Without the benefit of detailed evidence on systematic, subtle differences between the countries from the initial time point of our analysis (1970) to the late 1980s, it is difficult to see how such differences could have led to the dramatic divergence in married women's work patterns. In light of the cultural commonalities, an explanation giving primacy to patriarchal values would predict similar, not divergent, female labor force patterns. Indeed, the initial similarity in work patterns around 1970 suggests that shared cultural values once did lead to similar outcomes. In these earlier years, analysts of Taiwan noted that premarital work for one's parental family was followed by a postmarital retreat from the labor force (see Greenhalgh 1985; Jones 1984; Salaff 1981). Moreover, employers were eager to hire single women but reluctant to employ women after marriage (Diamond 1979; Kung 1983), illustrating the type of "marriage bar" phenomenon that existed at an earlier time in the United States, Britain, and Ireland (Goldin 1990; Pyle 1990). These patriarchal patterns, however, appear to have had much less impact on later Taiwan work patterns, which suggests that we must move beyond this level of explanation to account for differences between South Korea and Taiwan in the 1980s.

### Labor Demand Conditions

Both societies have been highly involved in the world market. As measured by export volume, however, Taiwan (with exports equal to 60% of GDP in 1986) has been more highly involved than South Korea (whose exports equal 35% of GDP; see Ranis 1992, p. 257; IBRD 1988, p. 243). Though neither society has had very much direct foreign investment of the sort that might skew capital intensity in production, South Korea has received massive foreign loans through subsidized government channels in ways that could skew production and labor markets.<sup>4</sup> These differences between the two societies have affected aggregate labor demand, educated labor demand, and job flexibility.

<sup>4</sup> To the extent that either society had direct foreign investment, some scholars conclude that the employment effects were either neutral or positive (Koo 1985; Ranis and Chi 1985; Tu 1990).

*Aggregate labor demand.*—One of the factors emphasized in export-led growth explanations is how government intervention in capital markets can distort the types of industries that are promoted. There are two principal ways that this could have an impact on the labor market for married women: (1) the relative emphasis on heavy versus light industry and (2) the emphasis on more capital intensive production, which leads to lower aggregate labor demand. We examine first the issue of industrial structure and its relationship to women's labor force participation (excluding agriculture) using 1989 labor force surveys for each country. Contrary to expectations, it is *not* the case that Taiwan's industrial distribution is more dominated than Korea's by light manufacturing or other industries that traditionally employ large numbers of women (see col. 1, appendix table A1). The industrial distributions of the two countries are in fact very similar, with some light industries such as food manufacturing, textiles, insurance, and real estate composing slightly more of the total nonagricultural labor force in Korea than in Taiwan. It is also striking that the proportion female in each industry is with few exceptions higher in Taiwan; across virtually all industrial categories, there are simply fewer women employed in Korea than in Taiwan.

While it appears unlikely that industrial structure per se is related to the lower employment rates of Korean women, it is much more probable that government policies encouraging capital intensity have had a significant effect through lowered labor demand and the concentration of "acceptable" white-collar jobs in large firms (we discuss the latter under *Job flexibility*, below). In the 1970s, South Korean production became more capital intensive while Taiwan production moved in the opposite direction. In South Korea a great influx of foreign capital has been funneled through government-subsidized loans to massive urban-centered business conglomerates (*chaebol*) that became more capital intensive (Amsden 1989; Hong 1979, 1983; Kim 1989; Mason et al. 1980). In contrast, Taiwan experienced little foreign borrowing and few government loans, and initial industrialization involved a myriad of small enterprises scattered throughout small towns and cities (e.g., Amsden 1991; Hamilton and Biggart 1988). By many accounts, the low levels of capital infusion and the highly labor-intensive methods in these small firms exhausted Taiwan's labor surplus as early as 1968 (Blank and Parish 1990; Galenson 1979; Lundberg 1979; Ranis 1992). Thus, the Korean government's preferential policies toward *chaebol* led to greater capital intensity while government policies in Taiwan led in the other direction, toward labor intensity. In short, the observed differences between the two societies are consistent with an export-led growth explanation that emphasizes the capital intensity of firms.

Changes in urban unemployment patterns over time are also consistent

with this mode of explanation. In Korea, urban unemployment continued to be a nagging problem for years (see table 3; see also Fields 1985; Galenson 1992). The large numbers of migrants who left the industry-impooverished countryside could not all find jobs in capital-intensive industry. In Taiwan, in contrast, urban unemployment remained at extremely low levels, and by the early to mid-1980s employers increasingly complained about labor shortages and the need to hire illegal foreign laborers (e.g., Speare, Liu, and Tsay 1988; Chang 1987). Similar trends in South Korea did emerge but not until the late 1980s, when unemployment began to be replaced by labor shortages, especially in small and medium-sized firms. This delayed labor shortage coincided with the belated rise in older Korean women's labor force participation in the late 1980s (see fig. 2). These parallels among the capital intensity of investment, the urban concentration of investment, urban unemployment rates, and married women's employment are consistent with the exported explanation.

*Educated labor demand.*—Unemployment of the well-educated is common throughout much of the developing world, and explanations of the phenomenon typically emphasize features such as “the diploma disease,” whereby educational credentials become the primary means to enter a modern sector that has dramatically higher incomes than the traditional sector (e.g., Dore 1976). The returns to education in South Korea as contrasted with Taiwan demonstrate the rampant credentialism in Korea. Heretofore, there has been a greater education premium in South Korea than in Taiwan (table 3).<sup>5</sup>

Regardless of its causes—be they the greater capital intensity of large firms or Korean employers' strong preference to hire the best graduates from the best schools—this pattern of educational returns has given parents and children a strong incentive to pressure the government to relax controls over secondary school and college enrollments. Consistent with this pressure, Korean male secondary school and university admissions grew dramatically in the late 1970s and early 1980s (see table 2 above). With the subsequent explosion in the number of secondary school and university graduates entering the labor market, it is perhaps not surprising that by the early 1980s unemployment among the educated was much greater in Korea than in Taiwan, where secondary school and university admissions remained more tightly controlled (table 3). As Michell suggests, “The Republic of Korea has a classic case of the diploma disease. . . . The imperfections of the labor market also mean that many well-educated people do not get jobs commensurate with their educa-

<sup>5</sup> As Amsden (1989) notes, the pattern of high returns to education in the midst of high unemployment among the educated is very curious and is difficult to explain.

TABLE 3  
LABOR MARKET CONDITIONS IN TAIWAN AND SOUTH KOREA

	1970	1975	1980	1985	1990
Unemployment rates (%).					
National:					
Taiwan .....	2.1	2.1	1.2	3.2	1.4
Korea .....	4.4	4.1	5.2	4.0	2.4
Urban:					
Taiwan (Taipei City) .....	NA	NA	1.7	3.6	1.8
Korea (Nonfarm) .....	7.4	6.6	7.5	4.9	2.9
Unemployment by education:					
Taiwan:					
Primary school .....	1.4	1.4	.5	1.5	1.3
Middle school .....	2.6	3.6	1.4	3.2	1.7
Secondary school .....	3.9	5.8	2.7	4.8	2.5
College .....	3.3	4.3	2.3	4.2	2.2
Korea:					
Primary school .....	NA	NA	2.7	1.5	0.7
Middle school .....	NA	NA	6.3	4.1	1.8
Secondary school .....	NA	NA	9.3	5.9	3.4
College .....	NA	NA	6.2	6.6	4.5
Paid-employee female/male income ratios:					
Taiwan .....	NA	51	.65	.65	.65
Korea .....	NA	42	.43	.47	.53
Paid-employee returns to education:					
Secondary/primary ratios:					
Taiwan .....	NA	1.47	1.23	1.19	1.12
Korea .....	NA	1.75	1.45	1.34	1.19
University/primary ratios:					
Taiwan .....	NA	2.36	1.94	2.14	1.79
Korea .....	NA	3.75	3.32	3.03	2.21
Employees in large firms*:					
Taiwan:					
500+ employees .....	NA	NA	8.6	7.4	7.7
100+ employees .....	NA	NA	26.9	26.5	23.9
Korea:					
200+/300+† .....	52.1	51.9	42.4	36.3	34.6

Sources — Taiwan, ROC (1987, 1990, 1991a, 1991b, 1991c), Korea, ROK (1991) and KSA (1991a, 1991b)

NOTE.—NA signifies that statistics were not available for this year.

\* This excludes government employment and includes in the denominator only firms with 10 or more employees

† This is 200+ employees through 1975, and 300+ employees starting in 1980.

tion" (1988, pp. 104–6). Korea produced too many college graduates in the 1980s for a society with only modest labor demand in the nonmanual sector (see Amsden 1989; Mason et al. 1980; McGinn et al. 1980). Moreover, when surveyed in 1984, Korean enterprises replied that their labor shortages were not in managerial, clerical, or sales jobs but in skilled manual (40%), unskilled manual (17%), and engineering jobs (14%–18%) (Amsden 1989, p. 225). In short, the queue of job applicants was likely to be the longest for educated applicants, with much of this applicant queue being composed of highly educated males.

From these observations on total labor demand and the long queues of highly educated males in South Korea, the low levels of Korean married women's paid employment begin to make sense. Much as in the United States or Britain before World War II, an adequate supply of educated males for clerical and other skilled jobs has provided little incentive for Korean employers to lower bars to married women's employment. This is in great contrast to Taiwan, where an inadequate supply of educated males has forced employers to alter their patriarchal preferences. As an example, commercial banks in Taiwan abolished the marriage bar in the late 1970s, and by the late 1980s private firms generally no longer asked women to quit upon marriage. In contrast, employer pressure on women to resign from white-collar jobs at the time of marriage has been very strong in Korea and has begun to diminish only in the late 1980s. With the passage of an Equal Opportunity Law in 1987, employer pressure on women to resign has become more indirect than previously (Korea Association of Women Friends 1989).

*Job flexibility.*—A final aspect of labor demand concerns the types of jobs common in the two societies. In South Korea, one of the consequences of government-subsidized loans to large urban businesses is that nonagricultural jobs are likely to be located in large firms in large cities, a situation that requires a lengthy journey to work and long, inflexible work hours (e.g., Amsden 1991; Deyo 1989; Hamilton and Biggart 1988; Koo 1985). In the 1970s, when many of the work habits of today were being established, over half of all Korean private-sector enterprises with 10 or more workers had more than 200 workers (see table 3). Even by 1990, if we define a "large" firm as one with 300 or more workers, over one-third of Korean private sector employees have continued to be in large enterprises. In contrast, less than 10% of Taiwan employees in private firms with 10 or more workers are in large firms of 500 or more. Even when one uses the very lax standard of 100 or more workers, only one-fourth of Taiwan firms are large. Thus, on average, the Taiwan employee works in enterprises of a more modest scale, where connections between employer and employee are likely to be more informal and arrangements for working mothers are probably easier to work out.



Commuting times have also remained longer in Korea. In 1990, the average Korean worker commuted 31 minutes to work, while even in Taiwan's largest city of Taipei, the average worker commuted only 24 minutes to work. In all large cities in Taiwan a worker averaged only 18 minutes, and in smaller cities he or she averaged only 13 minutes in commuting to work each day. Thus, for women, the tension between work and family appears to have been greater in Korea than in Taiwan. These facts fit both the export-led growth explanation, which emphasizes distortions due to government policy, and the labor supply explanation, which emphasizes the potential role of family responsibilities in keeping women out of the labor force.

In summary, the facts suggest that it is not just supply-side conditions or foreign-loans/government intervention that are important. Rather, it is the intersection of *supply conditions* (which are similar between the two societies) and the *demand conditions created by government policy and the nature of foreign loan investment* (dissimilar between the two societies). The development of capital-intensive production in South Korea kept aggregate labor demand lower than in Taiwan, while at the same time Korea's countryside was emptying migrants into the country's largest cities. The excess of educated workers was exacerbated by government policies that allowed the number of places at Korean universities to expand in response to the growing demand for education. Finally, foreign loans to the Korean government were allocated principally to *chaebol*, thus fueling the growth of large firms. Long commutes and long working hours have been the result. In all of these ways, the nature of labor demand has been shaped in such a way that the opportunities for married women—particularly the highly educated—to work in the formal sector have been hindered in South Korea to a degree that contrasts with the situation in Taiwan. We explore this in more detail with microlevel survey data.

#### MICROLEVEL ANALYSIS

If our macrolevel analysis is correct, the effects of women's supply-side characteristics should differ sharply in Taiwan and South Korea because these factors interact with very different demand-side conditions in the two societies. We first show how premarital work experience differently shapes postmarital work for women in Taiwan and South Korea; then we examine the different effects of supply-side factors on married women's current employment.

#### Work before and after Marriage

Examining young women's work before and after marriage shows how human capital accumulated through prior work experience promotes

## Married Women's Employment

TABLE 4

EMPLOYMENT STATUS AND OCCUPATIONAL DISTRIBUTION BEFORE AND AFTER  
MARRIAGE AMONG URBAN MARRIED WOMEN AGES 25-29\*

	TAIWAN				SOUTH KOREA			
	Before (%)	After (N)	Before (%)	After (N)	Before (%)	After (N)	Before (%)	After (N)
Self-employed .....	2.0	4	23.6	27	4.0	16	45.8	63
Family-enterprise worker .....	5.0	9	9.7	11	7.1	28	24.2	33
Employees .....	93.0	181	66.7	76	88.9	358	30.0	41
Administrative/managerial .....	.0	0	.9	1	.0	0	.0	0
Professional/technical .....	8.9	17	10.3	12	11.3	45	12.5	17
Clerical .....	32.0	62	31.1	35	37.3	151	1.4	2
Sales .....	3.9	8	3.7	4	3.4	14	.7	1
Service .....	10.2	20	8.4	10	2.0	8	4.6	6
Production .....	38.0	74	12.4	14	34.4	138	10.1	14
Farm .....	.0	0	0	0	.5	2	.7	1
Total in labor force .....	100.0	194	100.0	113	100.0	402	100.0	137
Not in labor force .....	5.8	12	45.1	93	20.6	104	72.9	369
N .....		206		206		506		506

SOURCES.—Taiwan Women and Family Survey (1989); Survey of Women's Employment (1985).

\* Educational levels of the two samples are standardized using the Taiwanese distribution as the reference.

work in paid (formal) employment in one society but not in the other. The sharp discontinuity between pre- and postmarital work gives further proof of the persistence of a marriage bar in South Korea (table 4).<sup>6</sup> The proportion of working women who are employees before marriage is very similar in the two societies (93% in Taiwan and 89% in South Korea). After marriage, the proportion of working married women who are employees drops to 67% in Taiwan; this contrasts with a more precipitous drop in South Korea to 30%. Conversely, the proportion of Korean working women in family enterprises and especially in self-employment rises sharply after marriage.

Turning to the occupational distribution among the formally employed, about one-third of single female workers in each society are clerical employees, another one-third are manufacturing employees, and the remainder are distributed throughout the major occupational categories. A higher proportion of single female workers in Taiwan are employees in service jobs, but other than that there are few premarital differences.

<sup>6</sup> In order to eliminate differences between the two countries' occupational distributions that might be produced by educational differences in the samples of women, we standardized Korea's educational distribution against Taiwan's for the purpose of table 4.

The occupational distribution of married women, however, is radically different in the two societies. In Taiwan, the proportion of all married women workers employed in clerical jobs is virtually *the same* as that for single women workers. But in South Korea, this figure drops from 37% *before marriage* to 1% *after marriage*. In short, it is extremely rare for young married female workers in South Korea to be in clerical jobs, whereas about one-third of this population of women in Taiwan are in clerical work. In contrast, the proportion of working women who are professional or technical employees remains about the same in the pre- and postmarital populations of Korea; this mainly represents women in the teaching profession, one of the few white-collar occupations open to married women.

One might ask whether married women eventually return to white-collar work. To investigate this, we also examined the premarital and current occupational distribution for women 45–49 years old (not shown). For this cohort also, the proportion of women workers in clerical jobs before marriage is similar in the two societies (around 20%). But about 14% of working women in Taiwan occupy such positions in their late forties, compared to 2% of women in South Korea. Thus our conclusions are not restricted only to young married women; older married women in South Korea also show an extremely low propensity to be involved in lower-level white-collar work. The proportions of older women in administrative/managerial work either before or after marriage are trivial in both societies, as is the case for the younger cohort as well. These comparisons further strengthen the conclusion that, for married women, the differences between the two societies in the quality of employment (specifically, access to white-collar jobs in the formal sector) are as great or greater than the differences in the propensity to be employed.

### Current Work

Analysis of whether married women currently work and, if so, in what types of jobs also demonstrates a very strong interaction between the characteristics a woman supplies to the labor market and the environment she encounters there.

*Methods.*—As independent variables, we emphasize the qualities a woman supplies to the labor market. The first group of qualities makes up her human capital. As already suggested, we infer that educational attainment will be positively related to women's formal employment in Taiwan but not in South Korea because of the greater competition from educated males and a continuing marriage bar in clerical jobs. We also include work experience in the equation and expect a similar result. Educational attainment is measured by four categories: primary school

or less, middle school, secondary school, or college or more.<sup>7</sup> Work experience is measured as whether the woman worked before marriage.

The second group of variables includes husband's income, measured as monthly earnings, and his employment status, which is categorized as employee, self-employed or employer, and not working.<sup>8</sup> We control for employment status because it is expected to have an important effect on the wife's type of labor force participation (e.g., wives of self-employed men are more likely to be family-enterprise workers). Our analytical focus is on the effect of husband's income *net* of this variable, with our expectation being that when labor market opportunities are constrained by a marriage bar the alternative of depending on husband's income will be more common.

The third group of variables represents household structure and family life-cycle characteristics (extended family structure, age of youngest child, number of children, and woman's age). With the exception of age of youngest child, these are included as control variables in the analysis. Household structure is considered extended if respondents lived with their parent(s) or parent(s)-in-law; it is considered nuclear otherwise. Age of youngest child is categorized into three groups: birth to two years, three to five years, and six years old or older. Woman's age is categorized as 25–29 years, 30–39 years, and 40–49 years. Distributions of the independent variables for both samples are shown in table 5.<sup>9</sup>

For the analysis of the determinants of married women's formal employment, we use a multinomial logit model that contrasts nonwork with formal employment, self-employed work, and family-enterprise work. We keep these three categories of work separate because of the major differences in the content and rewards of self-employment, family-enterprise work, and formal employment and because much of the implicit emphasis in our analytical models is on societal differences in formal employment. For each society we present two sets of models: a short model that includes women's human capital acquired before marriage

<sup>7</sup> While it would be desirable to retain junior colleges and universities as distinct categories, the number of women with university education is too low to permit this.

<sup>8</sup> We do this for two reasons: (1) classifying self-employed or family-enterprise workers by occupation is probably not conceptually meaningful and (2) our interest in the occupational distribution lies in the prediction of a flow of married women out of white-collar occupations. This is predicated on the idea of a marriage bar resulting from the actions of employers, and thus these actions are only relevant in cases where women are employees.

<sup>9</sup> For various reasons, married women ages 25–29 are underrepresented in the Taiwan sample. But because these are the women most likely to be in paid employment (see table 1 above), age biases in the sample serve not to overstate but instead to understate the differences we ascribe to Taiwan and South Korea. Moreover, aggregate statistics from government labor force surveys parallel the patterns we report here.

TABLE 5  
CHARACTERISTICS OF URBAN MARRIED WOMEN

	Taiwan	South Korea
Level of education:*		
Primary school or less .....	38.1	31.0
Middle school ..	16.0	28.5
Secondary school ..	30.4	33.3
College or more ..	15.5	7.2
Work before marriage.		
No .....	11.7	32.7
Yes .....	88.3	67.3
Average earnings of husband† .....	25,871	403,360
U S. dollar values ..	743.85	482.70
Husband's employment status:		
Employee .....	65.8	64.1
Employer/self-employed .....	31.1	26.5
Not working .....	3.1	9.4
Family type:		
Nuclear .....	84.9	88.3
Extended .....	15.1	11.7
Age of youngest child:		
0-2 years old .....	16.9	24.7
3-5 years old .....	19.9	24.5
6 or more years old .....	59.8	46.2
No children .....	3.4	4.6
Age.		
25-29 years old .....	12.3	29.7
30-39 years old .....	56.0	45.3
40-49 years old ..	31.7	25.0

SOURCES.—Taiwan Women and Family Survey (1989); Survey of Women's Employment (1985).

\* Values other than averages are percentages.

† Units are Taiwanese yuan and Korean won. U.S. dollar values are used for the analysis.

(education and work experience), husband's employment status, and family life-cycle characteristics, and a long model that adds husband's income.

*Results.*—The multinomial logit results for current types of work further demonstrate how women's employment in the two societies is shaped by conditions in the labor market (tables 6 and 7). As anticipated, education and premarital work experience affect married women's formal employment differently in the two societies. In the short model for Taiwan, education has a positive linear effect on a woman's probability of being an employee, with college education having a particularly strong effect. Experience in the labor market before marriage is also positively related to formal employment. The positive linear effect of education does not

hold for a woman's self-employment or for employment as a family-enterprise worker. Instead, these employment statuses are generally more likely to be occupied by the less educated.

In the short model for South Korea (table 7), education is negatively related to formal employment, although this effect is not linear: secondary school graduates are the least likely to work as employees, and the effect of college education is negative but not statistically significant. Work before marriage is not significantly related to formal employment after marriage. The general negative relationship between education and employment persists for self-employment and family-enterprise work as well, and work experience bears no relationship to these forms of employment. In both societies, wives whose husbands are employers or self-employed workers are less likely than other women to be employees and are much more likely to work in family enterprises (probably owned by their husbands).<sup>10</sup>

In addition to the direct effects of education, we anticipated that the link between Korean wives' education and formal employment would be mediated through their husband's income to a greater extent than in Taiwan. As shown in the long model for Taiwan (table 6), the coefficients for the effects of wife's education on formal employment remain quite stable when husband's income is included in the equation. Husband's income itself is negatively related to the probability of the wife working as an employee. In South Korea (table 7), the magnitude of the negative effect of husband's income on wife's formal employment is larger than in Taiwan. Moreover, controlling for husband's income has an impact on the influence of wife's education. Middle school education no longer has a significant dampening effect on wife's formal employment, the magnitude of the high school effect is reduced, and the effect of college education is positive rather than negative (although still statistically insignificant). This offers support for our assumption that a married woman's education has an indirect effect on her employment via husband's income in South Korea but not in Taiwan. In Korea an educated woman facing

<sup>10</sup> In models combining all three types of employment (not shown here), education has little effect on women's overall employment in Taiwan because of its offsetting effects on different types of employment. Only college education has a significant positive relationship with women's employment. In Korea, on the other hand, education has consistently negative effects on all three types of employment and thus shows a very strong negative relationship with aggregate employment. For both countries, the effect of the husband being an employer or a self-employed worker is also strongly related in a positive direction to the wife's working. This is because the positive effect on the wife's probability of being a family-enterprise worker (as well as self-employed in the case of Taiwan) overrides the negative effect on her probability of being an employee. These patterns hold in both the short and long models.

TABLE 6  
DETERMINANTS OF LABOR FORCE PARTICIPATION AMONG URBAN MARRIED WOMEN BY THEIR EMPLOYMENT STATUS: TAIWAN

	SHORT MODEL			LONG MODEL		
	Employee	Self-Employed	Family-Enterprise Worker	Employee	Self-Employed	Family-Enterprise Worker
Education:						
Primary school or less ..	-.16 (.21)	-.21 (.21)	-.25 (.31)	-.09 (.22)	-.17 (.21)	-.26 (.31)
Middle school ..						
Secondary school ..	.40* (.17)	-.92** (.21)	-.67* (.29)	.54** (.18)	-.83** (.22)	-.79* (.31)
College .....	1.51** (.21)	-.64* (.31)	-.59 (.45)	1.73** (.22)	-.50 (.32)	-.83 (.47)
Work before marriage:						
No .....						
Yes .....	.40* (.20)	1.04** (.25)	.98** (.37)	.39* (.20)	1.03** (.25)	1.01** (.38)
Husband's logged income .....				-.44** (.14)	-.33* (.14)	.65** (.24)
Husband's employment status:						
Employee ..						
Employer/self-employed .....	-.36* (.17)	.65** (.16)	2.71** (.27)	-.37* (.17)	.66** (.16)	2.63** (.27)
Not working .....	.16 (.36)	-.14 (.44)	-.20 (1.18)	-2.71** (.95)	-2.25* (1.02)	4.00* (1.96)

Family type:									
Nuclear .....	...	..	...	...	...	...	...	...	..
Extended .....	.53**	.07	.26	.49**	.04	.37	.04	.37	(.30)
	(18)	(.22)	(.30)	(.18)	(.22)	(.30)	(.22)	(.30)	
Age of youngest child:									
0-2 years old .....	...	.04	...	...	...	...	...	...	...
3-5 years old .....	39	(.29)	(.41)	.37	.04	.39	.04	.39	(.43)
6 or more years old .....	90**	59*	.20	.90**	58*	.25	58*	.25	(.44)
	(24)	(.29)	(.43)	(.24)	(.28)	(.44)	(.28)	(.44)	
No children ..	1.01**	-.24	-.62	.97**	-.29	-.43	-.29	-.43	(.98)
	(.37)	(.61)	(.96)	(.38)	(.61)	(.98)	(.61)	(.98)	
Age:									
25-29 years old ..	...	...	...	...	...	...	...	...	...
30-39 years old ..	-.39	.13	.72	-.35	.17	.66	.17	.66	(.43)
	(.22)	(.29)	(.42)	(.22)	(.29)	(.43)	(.29)	(.43)	
40-49 years old ..	-.42	-.09	.39	-.35	-.06	.41	-.06	.41	(.49)
	(.27)	(.33)	(.48)	(.27)	(.33)	(.49)	(.33)	(.49)	
Intercept ..	-1.38**	-2.08**	-4.44**	1.38	-.00	-8.70**	-.00	-8.70**	(1.66)
	(.31)	(.40)	(.62)	(.90)	(.98)	(1.66)	(.98)	(1.66)	
Observed proportion .....	31	18	.09	31	.18	.09	.18	.09	
Log likelihood .....	-1,859	-1,859			-1,841		-1,841		

SOURCES —Taiwan Women and Family Survey (1989), Survey of Women's Employment (1985)

NOTES —This table shows multinomial regression coefficients with each of the three types of work being compared to nonwork. Nos. in parentheses are SEs.  $N = 1,580$ .

\*  $P < .05$

\*\*  $P < .01$



TABLE 7  
DETERMINANTS OF LABOR FORCE PARTICIPATION AMONG URBAN MARRIED WOMEN BY THEIR EMPLOYMENT STATUS: SOUTH KOREA

	SHORT MODEL			LONG MODEL		
	Employee	Self-Employed	Family-Enterprise Worker	Employee	Self-Employed	Family-Enterprise Worker
Education:						
Primary school or less	...	...	...	...	...	...
Middle school	-.54** (19)	-.38* (18)	-.38 (26)	-.33 (20)	-.20 (18)	-.36 (26)
Secondary school	-1.32** (22)	-1.09** (19)	-1.24** (28)	-.91** (23)	-.72** (20)	-1.19** (30)
College	-.46 (30)	-.83** (32)	-1.46** (59)	.25 (33)	-.20 (34)	-1.37* (61)
Work before marriage:						
No	...	...	...	...	...	...
Yes	.22 (17)	-.02 (15)	-.25 (22)	.27 (17)	.03 (16)	-.28 (22)
Husband's logged income	...	...	...	-.86** (14)	-.80** (13)	-.07 (18)
Husband's employment status:						
Employee	...	...	...	...	...	...
Employer/self-employed	-.59** (22)	.03 (18)	5.58** (72)	-.61** (23)	.03 (18)	5.58** (72)
Not working	.32 (23)	.47* (22)	-6.02 (39,12)	-4.77** (85)	-4.26** (82)	-6.47 (39,20)

Family type									
Nuclear	...	...	...	...	...	...	...	...	...
Extended	...	...	...	...	...	...	...	...	...
	.06	.13	.08	.06	.04	.14	.06	.04	.14
	(.24)	(.30)	(.22)	(.25)	(.23)	(.31)	(.25)	(.23)	(.31)
Age of youngest child:									
0-2 years old	...	...	...	...	...	...	...	...	...
3-5 years old	...	...	...	...	...	...	...	...	...
	.97**	.59	.62**	.95**	.60**	.60	.95**	.60**	.60
	(.32)	(.35)	(.22)	(.33)	(.22)	(.35)	(.33)	(.22)	(.35)
6 or more years old	...	.58	.43	1.71**	.49	.55	1.71**	.49	.55
	1.65**	(.40)	(.26)	(.34)	(.26)	(.40)	(.34)	(.26)	(.40)
No children	...	1.86**	-1.31	1.42**	-1.42	-1.83**	1.42**	-1.42	-1.83**
	1.52**	(.64)	(.74)	(.41)	(.74)	(.64)	(.41)	(.74)	(.64)
Age:									
25-29 years old	...	...	...	...	...	...	...	...	...
30-39 years old	...	...	...	...	...	...	...	...	...
	16	-.42	.05	.24	.11	-.42	.24	.11	-.42
	(.27)	(.33)	(.21)	(.28)	(.21)	(.33)	(.28)	(.21)	(.33)
40-49 years old	...	...	...	...	...	...	...	...	...
	.08	-.34	-.14	14	-10	-.29	14	-10	-.29
	(.32)	(.42)	(.28)	(.32)	(.28)	(.42)	(.32)	(.28)	(.42)
Intercept	...	...	...	...	...	...	...	...	...
	-2.40**	-5.51**	-1.26**	2.44**	3.25**	-5.08**	2.44**	3.25**	-5.08**
	(.33)	(.79)	(.24)	(.84)	(.79)	(1.29)	(.84)	(.79)	(1.29)
Observed proportion	...	...	...	...	...	...	...	...	...
	.13	.09	.15	13	.15	.09	13	.15	.09
Log likelihood	...	...	...	...	...	...	...	...	...
	-1,486								

SOURCES.—Taiwan Women and Family Survey (1989), Survey of Woman's Employment (1985)

NOTES.—This table shows multinomial regression coefficients with each of the three types of work being compared to nonwork. Nos in parentheses are SEs  $N = 1,698$

\*  $P < .05$

\*\*  $P < .01$ .

the marriage bar in employment is more likely to use her educational credentials to compete in the marriage market for a husband who will provide an adequate income. Success in this marriage market will free her from working in the lower-status, lower-income jobs open to her.<sup>11</sup>

Based on our observation of aggregate work patterns by age and on information on the lower quality of jobs that are open to married women, we expected that Korean women with preschool children would be less likely to work than their counterparts in Taiwan. As predicted, although formal employment is positively related to the age of the youngest child in both societies, Korean women with a child under two years old are particularly likely to be nonemployees. The effects of youngest child's age are stronger for women in Korea than in Taiwan. This completes our list of labor supply conditions whose effects are shaped by the differing labor demand situation in each country.

Two general questions remain. First, are the differences in the effects of independent variables for the two societies statistically significant? Second, to what extent can the intersociety differences in married women's formal employment be attributed to the two societies' different distributions on the independent variables that measure women's human capital, husbands' characteristics, and household characteristics? To answer the first question, we tested for interaction effects between each independent variable and a dummy variable for the society. The difference in the effects of secondary school and college education on married women's formal employment are statistically significant for Taiwan and South Korea, and this holds whether the effects are direct or indirect (mediated through husband's income). The effect of husband's income is also significantly stronger in South Korea than Taiwan. Thus our major generalizations receive strong support. The difference in the effects of work experience and youngest child's age for the two societies is substantial but not statistically significant.

To answer the question of whether differences in the characteristics of Taiwanese and Korean women have a strong impact on the intersociety differences in employment probabilities, we calculated the probability of each type of employment based on the logit results given in tables 6 and 7 (substituting the mean for each independent variable); we also calculated the probability of employment in each society using the variable

<sup>11</sup> Although not shown here, the relationship between wife's education and husband's logged income, controlling for wife's age, is positive and strong in both countries. But the regression coefficient for wife's education is larger in South Korea than it is in Taiwan. The relationship between husband's and wife's education is also stronger in South Korea—which fits the social science literature describing how educated Korean women divert their energies from the labor market to the marriage market (e.g., Kim 1990; Park 1991).

## Married Women's Employment

TABLE 8

### ESTIMATED PROBABILITIES OF EMPLOYMENT AMONG URBAN MARRIED WOMEN BY EMPLOYMENT STATUS

	TAIWAN			SOUTH KOREA		
	Employee	Self- Employed	Family- Enterprise Worker	Employee	Self- Employed	Family- Enterprise Worker
Estimated directly from logit results .....	.309	.178	.042	.111	.159	.004
Estimated assuming sample composition of the other country	.311	.154	.023	.110	.128	.007

SOURCES.—Taiwan Women and Family Survey (1989), Survey of Women's Employment (1985)

means for the other society. Table 8 contains the results. The likelihood of Korean women working as paid employees is roughly one-third that for women in Taiwan, and these probabilities are barely altered when the sample composition of the other society is assumed. Also, consistent with the figures presented in table 5, the likelihood of self-employment and family-enterprise work is quite similar in the two societies. This is not strongly altered by standardizing each society for the sample characteristics of the other. Thus it is not the case that the differences we observe in urban married women's employment in Taiwan and South Korea are produced by differences in women's characteristics; rather, it is the way these characteristics are translated into employment that varies by society.

### DISCUSSION

Our analyses demonstrate the insufficiency of *labor supply* explanations in explaining married women's employment. Women's education, premarital work experience, and fertility are very similar in the two societies we examined. So, in aggregate comparisons between South Korea and Taiwan, labor supply conditions are a constant that cannot explain the different outcomes in the two places. Moreover, microlevel analyses show that labor supply conditions (as in the case of education and premarital work experience) lead to opposite outcomes or different magnitudes of outcomes (the case of child care responsibilities and husband's income) in the two societies. These results can be understood only by paying attention to the intersection between supply and demand factors in each society.

Much the same can be said for *patriarchal-value-based* explanations. In earlier periods, patriarchal-based explanations helped account for the very similar female labor market patterns in the two societies. Patriarchal culture supported a rigid marriage bar that kept young married women out of many jobs, particularly the better-paying, more secure jobs as paid employees. One can only understand why this marriage bar weakened in one society and remained strong in the other by paying attention to the intersection between cultural factors and demand factors. When total demand for labor remained relatively flaccid and the supply of educated males ample, employers found it easy to continue old cultural preferences for educated males. When demand for labor was strong and the supply of educated males inadequate, old cultural barriers fell. Thus, as in the United States and Britain after World War II, old cultural barriers to married women's work soon weakened. It is not that patriarchal-value-based explanations are wrong; rather they are insufficient without attention to the labor demand environment in which values are enacted.

*New international division of labor* accounts are inadequate for two reasons. First, early versions tend to lead to a uniform set of predictions for societies at the periphery of the world system. A foreign-investment-based prediction is not helpful, for direct foreign investment remained minimal in the two societies. When variation among societies is predicted by the new international division of labor, it is inconsistent with many of the observed patterns in Taiwan and South Korea. This view suggests that a society with greater world market involvement will have more labor market problems—a formal sector with menial or unskilled dead-end jobs open only to single women and a large informal sector made up of self-employed and family-enterprise jobs—but this is inconsistent with the two cases we have examined. Taiwan's exports totaled a staggering 60% of GDP compared to a figure of 35% for South Korea. Yet barriers to married women in the formal sector and the existence of a large informal sector more accurately characterize Korea than Taiwan.

An *export-led growth* framework proves more productive for explaining the divergence in married women's labor force participation in our two cases. Export-led growth explanations suggest how government distortion of factor markets can weaken otherwise positive effects of export-led growth on women's employment. This is consistent with the labor-intensive, dispersed industrialization pattern of Taiwan versus the more capital-intensive, large urban conglomerate industrialization pattern of South Korea. The Korean government's decision to promote large, capital-intensive enterprises by privileging them with foreign loans does appear to have had important consequences. The prospect of long commuting times and rigid working hours in large enterprises has created difficulties for urban Korean married women to balance work and family

responsibilities. Moreover, the government's promotion of large conglomerates in the 1970s raised wages for the more educated, which in turn caused parents to put more pressure on policymakers to relax secondary and higher education enrollment limits. This led to higher levels of educated unemployment and to longer job queues with males at the front and females at the rear. As a result, the fall of marriage bars to married females in white-collar jobs in the formal sector of the Korean economy has been considerably delayed in comparison to Taiwan.

We also draw two general conclusions from our analyses. First, we argue for the utility of reviving theoretical attention to the intersection of supply factors and demand factors in the labor market. Discussions of U.S. and British labor markets before World War II, the Irish labor market in the past few decades, and developing-country labor markets show that the analysis of married women's employment can be quite complex during periods of significant economic transition (Oppenheimer 1970; Standing 1981; Smock 1981; Goldin 1990; Mincer 1962; Smith and Ward 1984; Pyle 1990). Our results from Taiwan and South Korea demonstrate clearly that attention to the unique intersection of supply and demand is critical in understanding that complexity.

Our second general conclusion is related to the first. In order to understand properly the nature of labor demand in a country—a proximate determinant of women's labor force participation—it is important to pay attention to the ways that domestic government policy affects that demand. Accounts stemming from the "new international division of labor" tradition frequently fall short by predicting quite uniform consequences for societies at the periphery of the world market. Though export-led growth explanations are rather skeletal in providing only a checklist of market distortions to search for, they have the virtue of focusing attention on how the interaction between government policies and world market position shapes aggregate, urban, and educated labor demand. This is crucial for understanding the different labor market outcomes for married women in rapidly industrializing societies such as Taiwan and South Korea.

## APPENDIX

### Industry-Specific Demand

Appendix table A1 shows the surprising similarity in the industrial distribution of the labor force in Taiwan and South Korea. Moreover, there do not appear to be systematic industrial differences in the proportion of employed females in the two countries; in nearly all cases, there is a

TABLE A1

NONAGRICULTURAL INDUSTRIAL DISTRIBUTION AND WOMEN'S EMPLOYMENT SHARE  
IN TAIWAN AND SOUTH KOREA, 1989

	INDUSTRIAL DISTRIBUTION (%)			FEMALE SHARE (%)		
	Taiwan	Korea	(T - K)	Taiwan	Korea	(T - K)
<b>Primary Products:</b>						
Coal mining .....	.1	.3	-.2	11.1	5.4	5.7
Petroleum .....	1	.0	1	20.0	.0	20.0
Metal ore mining .....	0	1	-1	.0	.0	.0
Other mining .....	1	2	-0	10.0	10.0	0
<b>Light manufacturing:</b>						
Food .....	2.1	2.8	-.7	35.8	41.2	-5.4
Textile .....	7.3	9.2	-1.9	65.5	51.3	14.1
Wood .....	2.0	1.3	.7	27.9	19.0	8.8
Paper .....	1.9	1.8	.1	31.1	25.6	5.6
Chemicals, petroleum products ...	5.6	2.6	3.1	41.1	26.6	14.5
Nonmetallic .....	1.3	1.3	.0	34.7	20.5	14.3
<b>Heavy manufacturing</b>						
Basic metal .....	.9	.9	.0	15.9	6.2	9.7
Fabricating metal, machinery	15.4	11.3	4.0	36.2	22.5	13.8
Other manufactures .....	2.5	2.2	.3	53.1	45.8	7.3
<b>Public utilities:</b>						
Electricity, gas .....	.4	.3	.0	10.7	6.7	4.0
Water .....	.1	.1	.0	28.6	12.5	16.1
<b>Construction:</b>						
Construct-general .....	5.9	7.3	-1.4	13.0	7.2	5.8
Construct-special .....	2.8	1.7	1.1	6.0	7.8	-1.8
<b>Commerce:</b>						
Wholesale .....	5.0	3.2	1.8	61.6	29.1	32.5
Retail .....	13.2	18.0	-4.8	43.6	46.3	-2.7
Restaurants .....	4.2	7.2	-3.0	49.0	61.7	-12.7
<b>Transportation/business services.</b>						
Transport and storage .....	5.4	5.7	-.3	13.4	5.9	7.4
Communication .....	.8	.6	.2	24.6	21.8	2.8
Financial institutions .....	1.7	1.8	-.1	47.9	39.3	8.6
Insurance .....	.7	1.0	-.3	61.2	64.7	-3.4
Real estate, business services	1.9	2.7	-.8	42.9	21.3	21.6
<b>Services.</b>						
Public administration .....	4.3	3.4	.9	32.5	15.3	17.2
Sanitary services .....	.1	.2	-.1	40.0	19.2	20.8
Social services .....	7.7	6.8	.9	54.3	43.6	10.7
Recreational services .....	.8	1.2	-.4	45.5	34.6	10.8
Personal services .....	5.7	4.8	.9	48.9	54.7	-5.8
International bodies .....	.0	.1	-.1	0	11.8	-11.8
Average .....				39.0	34.6	4.4
Total .....	100.0	100.0	.0			

SOURCES —Taiwan ROC (1989), Korea ROK (1989)

NOTES —"T - K" = Taiwan % minus Korean % with a plus result meaning more in Taiwan and a minus result meaning more in Korea "Female share" = female as a % of all employees in each industry

higher proportion female in Taiwan.<sup>12</sup> When we standardize each country's industrial distribution to the other's, surprising results are produced (see Durand [1975] and Pyle [1990] for similar analyses for other countries). Korea's overall 34.6% female share *declines* to 32.2% after standardizing to Taiwan's industrial distribution, and Taiwan's 39.0% share *increases* to 40.7% when the industrial distribution is standardized to Korea's. This is the opposite of what would be expected if it were Korea's greater emphasis on heavy industry (or conversely, Taiwan's emphasis on light industry) that was producing country divergence in female employment. Only minor shifts occur when one standardizes on occupational rather than industrial categories, and the direction of change is once again in the opposite direction that one would predict. In short, across virtually all categories (other than a few limited categories such as restaurant and retail trade), be they heavy or light manufacturing, social services, or public administration, there are simply fewer women employed in Korea than there are in the comparable sector in Taiwan.

## REFERENCES

- Amsden, Alice H. 1989. *Asia's Next Giant. South Korea and Late Industrialization*. New York: Oxford University Press.
- . 1991. "Big Business and Urban Congestion in Taiwan: The Origins of Small Enterprise and Regionally Decentralized Industry (Respectively)." *World Development* 19 1121–35.
- Arnold, Fred, and Eddie C. Y. Kuo. 1984. "The Value of Daughters and Sons: A Comparative Study of the Gender Preferences of Parents." *Journal of Comparative Family Studies* 15 299–318.
- Beneria, Lourdes, and Shelley Feldman, eds. 1992. *Unequal Burden: Economic Crises, Persistent Poverty, and Women's Work*. Boulder, Colo.: Westview.
- Blank, Grant, and William Parish. 1990. "Rural Industry and Non-Farm Employment." In *Chinese Urban Reform*, edited by Reginald Kwok, William L. Parish, and Anthony Yeh. Armonk, N.Y.: M. E. Sharpe.
- Brinton, Mary C. 1993. *Women and the Economic Miracle: Gender and Work in Postwar Japan*. Berkeley and Los Angeles: University of California Press.
- Chang, Ching-hsi. 1982. "The Labor Force Participation of Married Women in Taiwan" (in Chinese). Pp. 97–127 in *The Chinese Economic Association Annual Conference Proceedings*. Taipei.
- . 1987. "Foreign Workers. A Preliminary Economic Analysis" (in Chinese). Pp. 198–216 in *The Fiftieth Symposium on Social Sciences*. Taipei: Academia Sinica, Institute of the Three Principles of Peoples.
- Chant, Sylvia. 1991. *Women and Survival in Mexican Cities*. Manchester: Manchester University Press.
- Chiang, L. N., and Y. Ku. 1985. *Past and Current Status of Women in Taiwan*. Taipei: National Taiwan University, Population Studies Center.
- Cho, Uhn, and Hagen Koo. 1983. "Economic Development and Women's Work in a Newly Industrializing Country: The Case of Korea." *Development and Change* 14:515–32.

<sup>12</sup> The percentage differences in this comparison are muted because the comparison includes unmarried as well as married women and because it includes women of all ages.



- Deyo, Frederic C. 1989. *Beneath the Miracle: Labor Subordination in the New Asian Industrialism*. Berkeley and Los Angeles: University of California Press.
- Diamond, Norma. 1979. "Women and Industry in Taiwan." *Modern China* 5. 317-40.
- Dore, Ronald. 1976. *The Diploma Disease*. Berkeley: University of California Press.
- Durand, John. 1975. *The Labor Force in Economic Development: A Comparison of International Census Data, 1946-1966*. Princeton, N.J.: Princeton University Press.
- Evans, Peter, and Michael Timberlake. 1980. "Dependence, Inequality, and the Growth of the Tertiary." *American Sociological Review* 45:531-52.
- Fields, Gary S. 1985. "Industrialization and Employment in Hong Kong, Korea, Singapore, and Taiwan." In *Foreign Trade and Investment: Economic Development in the Newly Industrializing Asian Countries*, edited by Walter Galenson. Madison: University of Wisconsin Press.
- Frobel, Folker, Jurgen Heinrichs, and Otto Kreye. 1980. *The New International Division of Labour. Structural Unemployment in Industrialised Countries and Industrialisation in Developing Countries*. Cambridge: Cambridge University Press.
- Galenson, Walter. 1979. "The Labor Force, Wages, and Living Standards." Pp. 384-447 in *Economic Growth and Structural Change in Taiwan*, edited by Walter Galenson. Ithaca, N.Y.: Cornell University Press.
- . 1992. *Labor and Economic Growth in Five Asian Countries: South Korea, Malaysia, Taiwan, Thailand and the Philippines*. New York: Praeger.
- Gereffi, Gary, and Donald L. Wyman, eds. 1990. *Manufacturing Miracles: Paths of Industrialization in Latin America and East Asia*. Princeton, N.J.: Princeton University Press.
- Goldin, Claudia Dale. 1990. *Understanding the Gender Gap*. New York: Oxford University Press.
- Greenhalgh, Susan. 1985. "Sexual Stratification: The Other Side of 'Growth with Equity' in East Asia." *Population and Development Review* 11:265-314.
- Hamilton, Gary, and Nicole Woolsey Biggart. 1988. "Market, Culture, and Authority: A Comparative Analysis of Management and Organization in the Far East." *American Journal of Sociology* 94:S52-S94.
- Hong, Won-tack. 1979. *Trade, Distortions and Employment Growth in Korea*. Seoul: Korea Development Institute.
- . 1983. "Export Promotion and Employment Growth in South Korea." In *Trade and Employment in Developing Countries*, vol. 1. Edited by Anne O. Krueger. Chicago: University of Chicago Press.
- IBRD (International Bank for Reconstruction and Development). 1988. *World Development Report 1988*. Washington, D.C.: Oxford University Press.
- Jones, Gavin W. 1984. "Economic Growth and Changing Female Employment Structures in the Cities of Southeast and East Asia." Pp. 17-60 in *Women in the Urban and Industrial Workforce: Southeast and East Asia*, edited by Gavin W. Jones. Honolulu: University of Hawaii Press.
- Kang, Seyoung. 1993. "Differences in the Process of Earnings Determination and Inequality between Women and Men in South Korea." *Korea Journal of Population and Development* 22:215-38.
- Kentor, Jeffrey. 1981. "Structural Determinants of Peripheral Urbanization: The Effects of International Dependence." *American Sociological Review* 46:201-11.
- Kim, Changsoo. 1989. "Labor Market Developments of Korea in Macroeconomic Perspectives." KDI Working Paper no. 8909. Korea Development Institute, Seoul.
- Kim, Kyung-Keun. 1990. "Schooling and Married Women's Work in a Developing Country: The Case of the Republic of Korea." Ph.D. dissertation. University of Chicago, Department of Education.
- Kong, Sae Kwon, and Minja Kim Choe. 1989. "Labor Force Participation of Married Women in Contemporary Korea." *Journal of Population and Health Studies* 9:116-36.
- Koo, Bohn Young. 1985. "The Role of Direct Foreign Investment in Korea's Recent

- Economic Growth." In *Foreign Trade and Investment: Economic Development in the Newly Industrializing Asian Countries*, edited by Walter Galenson. Madison: University of Wisconsin Press.
- Korea Association of Women Friends. 1989. *The Reality of Women Clerical Workers and Their Movement*. Seoul: Seok-Tap.
- KPHI (Korea Population and Health Institute). 1987. *Report of 1985 Fertility and Family Health Survey*.
- Krueger, Anne O. 1983. *Trade and Employment in Developing Countries: Synthesis and Conclusions*, vol. 3. Chicago: University of Chicago Press.
- . 1991b. *Social Indicators in Korea*.
- KSA (Korean Statistical Association). 1970, 1975, 1980, 1985, 1990, 1991a. *Annual Report on the Economically Active Population Survey*.
- Kung, Lydia. 1983. *Factory Women in Taiwan*. Ann Arbor, Mich.: UMI Research Press.
- KWDI (Korean Women's Development Institute). 1986. *Socioeconomic Indicators of Women's Position*.
- Lee, Kwang-kyu. 1985. "Development of the Korean Kinship System with Special Reference to Influence from China." *Bulletin of the Institute of Ethnology, Academia Sinica* 59:163–89.
- Lim, Linda Y. C. 1990. "Women's Work in Export Factories: The Politics of a Cause." Pp. 101–19 in *Persistent Inequalities: Women and World Development*, edited by Irene Tinker. New York: Oxford University Press.
- Liu, Paul K. C. 1984. "Trends in Female Labor Force Participation in Taiwan." Pp. 75–100 in *Women in the Urban and Industrial Workforce: Southeast and East Asia*, edited by Gavin W. Jones. Honolulu: University of Hawaii Press.
- London, Bruce. 1987. "Structural Determinants of Third World Urban Change." *American Sociological Review* 52:28–43.
- Lundberg, Erik. 1979. "Fiscal and Monetary Policies." Pp. 263–307 in *Economic Growth and Structural Change in Taiwan*, edited by Walter Galenson. Ithaca, N.Y.: Cornell University Press.
- Mason, Edward S., et al. 1980. *The Economic and Social Modernization of the Republic of Korea*. Cambridge, Mass.: Harvard University Press.
- McGinn, Noel F., et al. 1980. *Education and Development in Korea*. Cambridge, Mass.: Harvard University Press.
- Michell, Tony. 1988. *From a Developing to a Newly Industrialized Country: The Republic of Korea, 1961–82*. Geneva: International Labour Office.
- Mincer, Jacob. 1962. "Labor Force Participation of Married Women: A Study of Labor Supply." In *Aspects of Labor Economics*, edited by H. Gregg Lewis. Princeton, N.J.: Princeton University Press.
- Nash, June, and Maria Patricia Fernandez-Kelly, eds. 1983. *Women, Men, and the International Division of Labor*. Albany: State University of New York Press.
- Oppenheimer, Valerie Kincade. 1970. *The Female Labor Force in the United States*. Westport, Conn.: Greenwood.
- Papanek, Hanna. 1990. "To Each Less than She Needs, From Each More than She Can Do." Pp. 162–81 in *Persistent Inequalities*, edited by Irene Tinker. New York: Oxford University Press.
- Park, M. H. 1991. "Patterns and Trends of Educational Mating in Korea." *Korea Journal of Population and Development* 20.1–16.
- Pyle, Jean Larson. 1990. *The State and Women in the Economy*. Albany: State University of New York Press.
- Ranis, Gustav. 1979. "Industrial Development." Pp. 206–62 in *Economic Growth and Structural Change in Taiwan*, edited by Walter Galenson. Ithaca, N.Y.: Cornell University Press.
- . 1992. *Taiwan: From Developing to Mature Economy*. Boulder, Colo.: Westview.

- Ranis, Gustav, and Chi Schive. 1985. "Direct Foreign Investment in Taiwan's Development." In *Foreign Trade and Investment: Economic Development in the Newly Industrializing Asian Countries*, edited by Walter Galenson. Madison: University of Wisconsin Press.
- Reskin, Barbara F., and Patricia A. Roos. 1990. *Job Queues, Gender Queues*. Philadelphia: Temple University Press.
- ROC (Republic of China). 1966. *Statistical Abstract of the Republic of China*. Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- . 1987. *Yearbook of Manpower Statistics: Taiwan Area*. Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- . 1990. *Report on the Manpower Utilization Survey: Taiwan Area* (see also 1989, 1991). Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- . 1969, 1981, 1991a. *Monthly Bulletin of Manpower Statistics*. Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- . 1991b. *Social Indicators in Taiwan Area of the Republic of China*. Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- . 1991c. *Statistical Yearbook of the Republic of China*. Prepared by the Directorate General of Budget, Accounting, and Statistics. Taipei.
- ROK (Republic of Korea). 1970, 1975, 1980, 1985, 1990. *Population and Housing Census Report*. Prepared by the National Bureau of Statistics Economic Planning Board. Seoul.
- . 1972, 1977, 1982, 1987, 1991. *Statistical Yearbook of Education*. Prepared by the Ministry of Education. Seoul.
- . 1989. *Report on the Employment Structure Survey*. Prepared by the National Bureau of Statistics Economic Planning Board. Seoul.
- Salaff, Janet. 1981. *Working Daughters of Hong Kong*. New York: Cambridge University Press.
- Sassen, Saskia. 1988. *The Mobility of Labor and Capital: A Study in International Investment and Capital Flow*. Cambridge: Cambridge University Press.
- Smock, Audrey C. 1981. *Women's Education in Developing Countries*. New York: Praeger.
- Speare, Alden, Paul K. C. Liu, and Ching-lung Tsay. 1988. *Urbanization and Development*. Boulder, Colo.: Westview.
- Standing, Guy. 1981. *Labour Force Participation and Development*, 2d ed. Geneva: International Labour Office.
- Thurrow, Lester. 1969. *Poverty and Discrimination*. Washington, D.C.: Brookings Institution.
- Timberlake, Michael, ed. 1985. *Urbanization in the World-Economy*. Orlando, Fla.: Academic Press.
- Tinker, Irene, ed. 1990. *Persistent Inequalities*. New York: Oxford University Press.
- Tu, Jenn-hwa. 1990. *Direct Foreign Investment and Economic Growth: A Case Study of Taiwan*. Monograph Series no. 48. Taipei: Academia Sinica, Institute of Economics.
- Ward, Kathryn, ed. 1990. *Women Workers and Global Restructuring*. Ithaca, N.Y.: ILR Press.
- Wolf, Diane L. 1990. "Linking Women's Labor with the Global Economy." In *Women Workers and Global Restructuring*, edited by Kathryn Ward. Ithaca, N.Y.: ILR Press.

# 'Til Death Do Us Part: Marital Disruption and Mortality<sup>1</sup>

Lee A. Lillard  
*RAND*

Linda J. Waite  
*University of Chicago*

Both men and women appear to benefit from being married. This article uses data from the Panel Study of Income Dynamics to examine the extent to which three key factors—financial well-being, living arrangements, and marital history—account for this relationship. The authors model mortality using a flexible hazard model and find that both married men and women show substantially lower risks of dying than those who are not married. The study's results suggest that—for women but not for men—the improved financial well-being that often accompanies marriage accounts for much of its beneficial effect. For both husbands and wives the benefits from marriage appear to cumulate as the length of the union increases.

## INTRODUCTION

Marriage appears to protect its incumbents against many of life's blows. The married, especially men, show lower levels of alcohol and cigarette consumption, higher earnings, and, perhaps as a consequence, lower levels of mortality than the unmarried. The relationship between marriage and death rates has now reached the status of a truism, having been observed across numerous societies and among various social and demographic groups. These studies also generally agree that the benefits of marriage over singlehood are larger for men than for women (Gove 1973).

At least two processes could lead to the result that married persons

<sup>1</sup> This research was supported by grant no. P01-AG08291 from the National Institute on Aging. We want to recognize Mary Layne, who carefully prepared the complex marital and mortality histories. The software for this application was designed and written by Lee A. Lillard and Constantijn W. A. Panis. Please address all correspondence to Linda J. Waite, Population Research Center, 1155 E. 60th Street, Chicago, Illinois 60637.

live longer than single persons. First, marriage may be more available to those in good health than to those in poor health, so that this selectivity results in higher mortality rates for those not married. Second, marriage itself could cause lower mortality rates, presumably by increasing the health of married persons relative to their unmarried peers and by reducing risky behaviors.

This article focuses primarily on the second of these mechanisms—the protective effect that marriage has on the partners. It uses data from the Panel Study of Income Dynamics (PSID) to follow men and women into and out of marriages, comparing risks of dying by marital status and duration of marriage. We model mortality using a flexible hazard model in which the risk of dying can change with changes in marital status and with marital duration. We examine the extent to which three key factors account for the mortality advantage of marriage: financial well-being, living arrangements, and marital history.

## BACKGROUND

Most previous analyses of the link between marriage and mortality have focused on cross-sectional differences between death rates for the married and the unmarried as evidence of a connection. These studies often use information from a sample of death certificates, which give demographic information on the decedent, together with population figures from other sources, such as the Current Population Survey or decennial census, for appropriate denominators with which to calculate death rates (Gove 1973; Kobrin and Hendershot 1977; Trovato and Lauris 1989). Alternatively, studies focus on a measure of relative risk, the relative mortality ratio, which compares the death rate for the married group to the death rate for the those who are unmarried (Goldman 1993). These studies consistently show mortality advantages for the married compared to the unmarried.

The use of information from death certificates has some strengths; it allows analysis of the exact cause of death, since this information is contained on death certificates. And the related approach, which requires only information on the number of deaths by marital status and the size of the married and unmarried populations, can be applied to a large number of countries and time periods. These approaches have two important limitations, however. First, they greatly constrain the characteristics of the individuals available for analysis, often to only age, sex, race, and marital status. Second, the cross-sectional data do not allow researchers to follow individuals over time, into and out of various marriages. Although these data contain the marital status of the person at the time of death, they tell us nothing about marital history, recent marital transi-

tions, or the duration in various marital statuses. In addition, this approach provides no information about the living arrangements of the person, the person's financial situation, or other key measures of the person's circumstances.

Some studies have used a partial panel approach, collecting baseline information on demographic and health characteristics of a sample of individuals, then tracking their deaths over an extended period (Kotler and Wingard 1989; Rosengren, Wedel, and Wilhelmsen 1989). These studies also find significant differences between death rates of the married and unmarried, even when differences in characteristics and initial health behaviors are taken into account. However, such studies can consider only characteristics of the person at the beginning of the period. Changes over time in marital status, living arrangements, or financial well-being cannot be traced or modeled. Thus, those who divorce or become widowed are grouped with those who remain married in these analyses. These analyses cannot examine the effect of transitions into and out of marriages or the effect of marital history on the risk of dying. Mott and Haurin (1985) apply a variant of this approach, using characteristics of decedents in the year before they die and characteristics of survivors in randomly assigned years to predict mortality. They find that being married strongly affects survival for older men and that numerous controls diminish only slightly the size of these effects.

Several recent studies have specifically examined transitions into and out of marriage, and their findings are informative. Helsing, Szklo, and Comstock (1981) use baseline data from a 1963 health survey in a local area together with information collected in 1975 on later changes in marital status and mortality during the period. They focus on the transition from married to widowed and from widowed to remarried and find that men who were recently widowed face substantially higher death rates than either men who had never married or those widowers who remarried. They found the same relationship, although somewhat attenuated, among those who survived at least three years after a wife's death. The authors argue that both selection and protective effects were operating; the healthier widowers were more likely to remarry, and remarriage appeared to protect them from the stresses of the previous loss.

Zick and Smith (1991) use longitudinal data from the PSID to examine the effects of recent marital transitions on mortality. They identify those who had married, divorced, or become widowed in the two previous years and estimate the effects of these changes on the hazard of dying, net of marital status in the previous year. This study focuses especially on how poverty spells in the three preceding years affect mortality. It also tests the hypothesis that mortality risks rise following any transition out of marriage, because these transitions cause high levels of emotional

distress. Zick and Smith find that marital status and recent marital transitions matter for men but not for women. Those men who have been divorced or widowed for three years or more face a significantly greater risk of mortality than married men do. In addition, they find that recent divorce (but not widowhood) increases men's mortality risk.

Although these recent studies examine the effect of changes in marital status on risk of dying, they use relatively simple measures of these transitions rather than looking at the full range of movements into and out of marriages, and they use very limited measures of duration in various marital statuses. The research reported here moves well beyond these earlier analyses to examine the effects of (1) movements into and out of marriage; (2) transitions from married to three different unmarried statuses—separated, divorced, and widowed; and (3) marriage duration. We explicitly compare the currently married with four categories of the unmarried: never married, separated, divorced, and widowed. We use a flexible hazard model to model movements into and out of marriages and the impact of marital duration.

#### THEORETICAL FRAMEWORK

Scholars interested in explaining the observed relationship between marriage and mortality have pointed to two distinct effects: a selection effect, by which those least likely to die are the most likely to marry, and a protective effect, whereby marriage itself decreases the chances of dying. The workings of the first mechanism seem straightforward; those with chronic conditions or dangerous or unhealthy lifestyles may have more trouble attracting a spouse than healthy, relatively settled individuals (Carter and Glick 1976; Retherford 1975). Goldman (1993, p. 190) points out that "spouses may be selected for better health not only through the direct exclusion of mentally and physically ill persons from marriage, but also through a wide range of selection criteria including income, physical appearance, risk-taking behavior, health-related habits such as smoking and excessive drinking and emotional stability."

The protective effects of marriage may be more complex, and we focus on these in this research. Kobrin and Hendershot (1977) argue that the social integration provided by marriage results in lower mortality. They find that persons with high-status social ties—men or women heading families—have the lowest levels of mortality in the cross section, whereas those with few social ties—men living alone—or low-status ties—women living as dependents in families—have the highest mortality. (But Kotler and Wingard [1989] find that increasing numbers of children have no effect on mortality for employed women and a *positive* relationship with

mortality for housewives.) According to Pearlin and Johnson (1977, p. 714), "Marriage does not prevent economic and social problems from invading life, but it apparently can help people fend off the psychological assaults that such problems otherwise create."

Trovato and Lauris (1989) argue that mechanisms other than social integration may operate to lower the mortality of the married. First, the married tend to have higher levels of economic resources than the unmarried. Married women benefit financially from the generally higher incomes of their husbands, and married men benefit financially from the economies of scale associated with marriage and from the household management skills of their wives.

Second, marriage may encourage healthy behaviors and discourage risky or unhealthy ones. Umberson (1987) finds that marriage and parenthood reduce the occurrence of a number of health-threatening behaviors such as problem drinking, drinking and driving, substance abuse, and other risk taking.

#### SEX DIFFERENCES IN BENEFITS OF MARRIAGE

The literature on the relationship between mortality and marriage consistently finds that marriage benefits men more than women, at least to the extent that the difference between the death rates of the married and unmarried is substantially greater for men than for women (Gove 1973; Trovato and Lauris 1989; Hu and Goldman 1990). This suggests that either the selective or protective effects of marriage, or both, operate more strongly for males than for females.

Why should marriage offer more protection against dying to men than to women? A number of possibilities have been suggested. First, single men, especially the divorced and separated, seem to be especially prone to risky behavior and unhealthy lifestyles—much more than single women, and married men are especially ill-disposed to such behavior (Umberson 1987). So risk taking and its connection with nonmarriage for men much more than for women could explain the relationship. A second explanation focuses on sex differences in social integration; to the extent that women take responsibility for maintaining relations with family and friends, unmarried men may find themselves relatively isolated, whereas unmarried women may be embedded in a network of social support. Married men receive the benefits of social integration and support from their wives and through her from others. Litwak and Messeri (1989) suggest that women's social networks often include members of the medical community, allowing them to manage the nontechnical aspects of health care within the family. Married men also appear to benefit directly



from a wife's usually substantial contribution to household labor and from the organized and regularized family life that tends to flow from it (Umberson, Wortman, and Kessler 1992).

## HYPOTHESES

In this section we develop a series of hypotheses about the relationship between marriage, and between transitions into and out of marriage, and risks of dying. In our model, exposure to risk of mortality depends on the individual's own resources and those of their spouse, their living arrangements, and their gender. It also depends on their current marital status and marital history.

As we have already stated, the married may face lower risks of dying because those least likely to die are selected into marriage. In addition, marriage itself offers protection against mortality. Note that, as Helsing et al. (1981) argue, both processes could operate at the same time, with some selection of those in good physical, emotional, and mental health into marriage and some protective effects conferred by wedlock. Although we do not directly address the existence of positive selection of the married, we try to point out in our discussion the mechanisms through which such selection could influence the results we observe.

The protective effect of marriage could result from an increase in material well-being stemming from the economies of scale achieved by combining resources and from the specialization of tasks common in marriage (Becker 1981). Marriage could protect individuals by encouraging healthy behaviors, such as regular meals and hours, and by discouraging unhealthy ones, such as alcohol consumption or smoking (National Center for Health Statistics 1988). Finally, some scholars have argued that marriage protects individuals by integrating them socially with others—into the family, the community, the schools, and religious institutions.

We can test some of this reasoning directly. First, using measures of household income, we can measure the improvement in material well-being associated with marriage. To the extent marriage improves health and reduces mortality by improving material well-being, holding constant the financial resources of the household should reduce the effect of marital status *per se*.<sup>2</sup>

Second, married individuals are integrated socially primarily through their connection to their spouse and secondarily through their connection to others with whom they live and through whom they are connected to family and friends outside the household. Unmarried individuals who

<sup>2</sup> Theories dealing with the impact of income on health generally do not distinguish between the subject's own income and household income.

live with others—their own children, their parents, other relatives, roommates, or a cohabitor—may receive many of the benefits of the social integration that comes from marriage (Kobrin and Hendershot 1977). We test this idea directly by measuring the number of adults (other than the spouse) and the number of dependent children (under age 18) who share the household with the individual of interest. If some of the benefit of marriage results from living with others, we will observe a decrease in the mortality differential when we hold constant the presence of other adults and children.

This reasoning also suggests that the effect of living with an adult (other than the spouse) is larger for the unmarried than for the married. This would be the case (1) because of the longevity and strength of the tie to a spouse compared to ties to another person and (2) because married persons have a close tie to another person in the household not available to the unmarried living alone, so the marginal benefit to the married of another adult is smaller than the marginal benefit to a single adult living alone.<sup>3</sup> Of course, those unmarried persons who live with another adult may be selected—either positively or negatively—from the pool of those currently not married. Those least able to manage on their own may live with parents or other relatives, so that selection is negative. Positive selection might mean that those with the most to offer a potential mate marry, those with somewhat less to offer cohabit, and those with the least to offer remain single and living alone. This reasoning fits with Booth and Johnson's (1988) conclusion that cohabitators tend to have personal characteristics that make them poor marriage material.

A key feature of our analysis is the use of a longitudinal data set, in which many individuals enter and leave marriage. Because we thus observe many of the same individuals in various marital statuses, we are able to estimate the effect of changes in marital status on the risk of dying. To the extent that the relationship between marriage and mortality results from the selection of the healthy into marriage, we would expect to see a reduction in the hazard of dying with marriage since marriage would serve to identify healthy individuals. The same argument applies to the selection into marriage of those who are least prone to risky behavior.<sup>4</sup> To the extent that marriage reduces the prevalence of risky behav-

<sup>3</sup> Much social integration comes from ties to those outside the household. However, we argue here that those with spouses have more potential sources of ties to others outside the household than the unmarried, since the married may have ties both on their own and through their spouses. In addition, the married have less need of ties to those outside the household, since they can meet more of their own needs through spouses than can the unmarried living alone.

<sup>4</sup> We would like to hold constant the health and risk-taking behavior of individuals prior to marriage. We are unable to do so, however, because we observe a large proportion of our sample a number of years after marriage, and because the PSID

iors like drinking and driving, substance abuse, and binge drinking—especially by men—then we may see an immediate effect of marriage on the hazard of dying from accidents or violence.

We also hypothesize that protective effects of marriage confer benefits over an extended period of time. If marriage reduces risky behaviors and increases material well-being, the effects of this change on mortality may appear over time, as the individual reaps the benefits of better health. This reasoning suggests that the mortality advantage of being married increases with length of marriage up to some point.<sup>5</sup>

A large body of research (Brody 1985; Antonucci and Akiyama 1987) points to women's greater integration than men in networks of both kin and friends. Women tend to do the personal and emotional work of maintaining family connections with extended kin and with friends (Di Leonardo 1987). Women also tend to perform most housekeeping tasks within the home (Goldscheider and Waite 1991). At the same time, men have significantly higher earnings, on average, than women over their lifetimes and substantially more access to pension income after retirement (Holden, Burkhauser, and Feaster 1988).

These gender differences in adult roles mean that men and women often have very different experiences after widowhood or divorce. Since men tend to cede network maintenance to women, single men often find themselves without much social support. And since men have left much of the homemaking to women, widowed and divorced men often find that their meals are irregular and their homes unkempt, whereas newly single women more often find themselves short of money (Hoffman and Duncan 1988). To the extent that the protective effect of marriage derives from social integration and a settled home—both of which should encourage healthy behaviors—men should be more disadvantaged from losing a spouse than are women.<sup>6</sup> This reasoning suggests that holding constant the living arrangements of the individual will reduce the male/female

---

lacks repeated measures of health status. The PSID also lacks annual measures of cigarette and alcohol consumption.

<sup>5</sup> One reviewer pointed out that, even if marriage does not change behavior, those married persons with the worst health and riskiest behavior tend to leave the population through death or divorce, leaving in the married population those with the best health and lowest levels of risk taking. However, our model implicitly compares those married at a particular duration to those unmarried at the same age and time period, who presumably face even stronger selection out of the population through death than their married counterparts. Goldman (1993) points out that selectivity into marriage by health status does not have simple or intuitively obvious effects on the relative death rates of single and married persons.

<sup>6</sup> Previously married men under age 65 show the highest rates of alcohol consumption of any group (National Center for Health Statistics 1988).

differential in the impact of the end of marriage on mortality and that holding constant household income will also decrease this differential.

### DATA

The data used in this analysis come from the PSID, a large, national longitudinal data set that began in 1968 with 5,500 households. The sample has been resurveyed each year since that time.<sup>7</sup> In each year the PSID obtained a detailed interview with the household head on his or her demographic and socioeconomic characteristics and those of household members. The PSID contains detailed annual information on marital status and a retrospective marital and fertility history of the respondent and spouse, which was obtained in 1985. This retrospective history, together with annual information on marital status, allows us to precisely date all marriages, separations, and divorces.

The PSID allows us to follow a representative sample of married couples over a lengthy period during which some die. We can observe changes in marital status over time and the responses to those changes. We can hold constant characteristics of the individuals and their marriage and compare their experiences to those of the unmarried.

Every individual of any age living in a PSID sample household in 1968 has been followed since that time. Thus, we observe those who were children in the early years of the survey as they reach adulthood, marry, and divorce, and we observe the later experiences of a large number of couples who were married at the first survey. The analysis presented here uses the augmented file of the PSID, which includes information on all sample members ever interviewed in the survey, including those who died. We verified dates of death for a majority of decedents using information from death certificates provided to the PSID by states.<sup>8</sup>

The sample used here consists of 11,112 individuals, 5,053 men, and 6,059 women. We observe all sample members 10 years old or older at any point between 1968 and 1985 and follow them through the entire 17-year period.<sup>9</sup> During that period, among those married, we observe 706 individuals who become widowed at least once and 857 individuals who die.

<sup>7</sup> Each member of an original sample household, or any child born to a sample member, is followed whenever he or she exits a sample household. New households formed by sample members are followed as new sample units and interviewed in the same way as original households. Thus, the sample has grown over time.

<sup>8</sup> We have the exact date of death from death certificates for 68% of all decedents.

<sup>9</sup> We include those persons 10 years old and older to ensure observation of early marriage and as a baseline for older ages.

TABLE 1  
NUMBER OF INTERVALS UNMARRIED AND MARRIED, 1968-85

	MEN		WOMEN	
	Respondents	%	Respondents	%
Intervals unmarried:				
0 . . . . .	1,969	39.0	1,782	29.4
1 . . . . .	2,581	51.1	3,660	60.4
2 . . . . .	476	9.4	576	9.5
3 . . . . .	27	.5	40	.7
Intervals married:				
0 . . . . .	869	17.2	1,608	26.5
1 . . . . .	3,840	76.0	4,119	68.0
2 . . . . .	328	6.5	315	5.2
3 . . . . .	16	.3	17	.3

Tables 1 and 2 below present information on marriage, death, widowhood, and divorce for males and females in the sample. Table 1 shows the number of intervals during which we observe individuals as either married or unmarried. One individual can be observed first as never married, then as married, then as separated, then divorced, then as (re)-married, then as widowed. In this example we would observe six separate intervals for this one person. In this table individuals who have more than one "married" interval ended the original marriage—through widowhood, separation, or divorce—and entered a later marriage. The same is true of individuals who have more than one unmarried interval; they must have entered and left at least one marriage. Table 1 shows that 39% of the men in our sample were never observed in the unmarried state between 1968 and 1985, and 51% were unmarried for only one interval; during this interval they were either never married, separated, divorced, or widowed. Three-quarters of the men and two-thirds of the women were observed in the married state exactly once; if that marriage ended, then they did not remarry—and the marriage may have ended with their death. Table 1 shows that relatively few men or women were observed in two or more marriages in the period.

Table 2 shows the experience of persons from 1968 through 1985. This table shows the percentage experiencing each of the transitions—divorce, widowhood, and death—at any point in the period; some individuals experienced more than one type of transition. Of those married at the first interview, note that more than twice as many of the men as women die during the period, with women much more likely than men to become widowed. The similarity of the percentages divorced reflects the fact that

TABLE 2  
TRANSITIONS EXPERIENCED, 1968-85

GENDER	N	MARRIED IN 1968			UNMARRIED IN 1968		
		% Divorced	% Widowed	% Deceased	N	% Married	% Deceased
Men .....	2,627	17.7	7.4	14.4	2,426	64.2	4.3
Women .....	2,658	16.7	16.2	6.6	3,401	52.7	5.9

the men and women who were married in 1968 were married to each other. Of those initially unmarried, most married and very few died by 1985. The percentage marrying was higher for men than for women, but the percentage dying was slightly higher for women, due in part to the overrepresentation of older widows among unmarried women. The lower percentage dying among those unmarried in 1968 than those married primarily reflects the much younger average age of the unmarried.

#### MODEL

Our model is designed to study the effects of marital status and marital transitions on the hazard of mortality over the 20-year period covered by the PSID, 1968-88, and to control for the effects of other relevant covariates that might account for the gross difference. We use a continuous time hazard or failure-time framework.

The mortality hazard rate at age  $a$  is the instantaneous probability of dying at age  $a$  conditional on having survived to that age. We use a generalized Gompertz formulation in which the log hazard,  $\ln h(a)$ , is a linear function of age, calendar time, marriage duration, marital status, and other covariates:

$$\ln h(a) = \alpha_0 + \alpha_1' \text{Age}(a) + \alpha_2' \text{Time}(a) + \alpha_3' \text{MarDur}(a) + \alpha_4' \text{UnMar}(a) + \alpha_5' X(a),$$

where the explanatory variables and their definitions are given in Table 3.

The hazard equation may be written as the product of (1) a "baseline" hazard that is a linear function of age and generates duration dependence in the hazard and (2) a term generating proportional shifts of the baseline hazard due to the effects of covariates that are either constant over time or that are time varying but constant within subintervals of time. That is,

$$h(a) = e^{\alpha_0 + \alpha_1' \text{Age}(a) + \alpha_2' \text{Time}(a) + \alpha_3' \text{MarDur}(a)} e^{\alpha_4' \text{UnMar}(a) + \alpha_5' X(a)}.$$

TABLE 3  
DEFINITIONS OF VARIABLES

Variable	Description
Age( <i>a</i> )	Piecewise linear spline with nodes at 40, 60, and 75 years (begins at age 10)
Time( <i>a</i> )	Calendar time trend, 1968–85
Marital status:	
MarDur( <i>a</i> )	Years since beginning current marriage (zero if not married)
UnMar( <i>a</i> ):	
Unmarried	Indicator for unmarried (vs. married)*
Separated	Indicator for separated (vs. never married)
Divorced	Indicator for divorced (vs. never married)
Widowed	Indicator for widowed (vs. never married)
Other regressors, <i>X</i> ( <i>a</i> ):	
Household income	Log of household income in 1980 dollars; linear spline with nodes at quartiles of the income distribution†
Other adults in household	<i>N</i> of adults (other than spouse) over age 18 in the household
Dependent children in household	<i>N</i> of children under age 18 in the household
Black	Indicator for race
Education < 12	Less than 12 years of schooling
Education ≥ 16	16 or more years of schooling
Spouse's education < 12	Spouse has less than 12 years of schooling
Spouse's education ≥ 16	Spouse has 16 or more years of schooling
Northeast region	Current region of residence
Large city	Residence in a city with a population of 100,000–499,000
Very large city	Residence in a city with a population of 500,000 or more

\* At zero marriage duration, where relevant.

† Quartile cutpoints: 8 5, 9 6, 10 7

The baseline hazard is a function of age, calendar time, and marriage duration. Age(*a*) is a piecewise linear spline with changes of slope at ages 40, 60, and 75. Time(*a*) is a linear trend within the 1968–88 period. MarDur(*a*) is duration of the current marriage at age *a* if married, but zero if unmarried. A person's age and calendar time move together collinearly, but their separate effects are identified, since persons are born at different points in time. For a married person, marriage duration is also collinear with age and time, so the combined duration dependence is given by the sum of the three effects.

Covariates that shift the baseline hazard proportionately include (1) an

indicator variable for unmarried (relative to married) and separate indicator variables for separated, divorced, and widowed (each relative to never married), all in the vector  $\text{UnMar}(a)$ , and (2) all other covariates,  $X(a)$ , that may vary with age. The vector  $X(a)$  includes the demographic characteristics of race (an indicator for black), education, education of spouse if married, northeast region, residence in a large or very large city, household income (a piecewise linear spline in the log of household income with slope changes at income quartiles), and living arrangements (the number of adults<sup>10</sup> and the number of children<sup>11</sup> living in the household).

The "baseline" survivor function,

$$S_0(a) = \exp \left\{ - \int_{a_0}^a \exp[\alpha_0 + \alpha_1' \text{Time}(\vartheta) + \alpha_2' \text{Age}(\vartheta) + \alpha_3' \text{MarDur}(\vartheta)] d\vartheta \right\},$$

incorporates all sources of duration dependence and is defined from age 10 ( $a_0 = 10$ ). The corresponding survivor function gives the probability of surviving to at least age  $a$  conditional on surviving to initial age  $a_0$  and incorporates the effects of all other covariates. The survivor function is given by

$$S[a|a_0, H(a_0, a)] = \prod_{i=1}^I \left[ \frac{S_0(a_{i+1})}{S_0(a_i)} \right]^{\exp[\alpha_4' \text{UnMar}(a_i) + \alpha_5' X(a_i)]},$$

where  $H(a_0, a)$  denotes the history of the covariates ( $\text{UnMar}[t]$  and  $X[t]$ ) from  $a_0$  to  $a$ . The subintervals for which the time-varying covariates are constant define  $i$  ( $i = 1, \dots, I$ ).<sup>12</sup>

Estimation is based on maximum likelihood. For some persons, we observe the date of death, but for others we know only that the person is still alive as of the last survey date. We define  $a^*$  to be either date, as appropriate, and the indicator variable  $D(a^*)$  to be "1" if it is the date of death and "0" otherwise. Every person in the PSID sample was alive at the initial 1968 survey date; we denote age at that time by  $a_{68}$ . The likelihood function is given by the probability of the observed outcome (death date or censor date) conditional on having survived until 1968:

$$L[a^*|a_{68}, H(a_{68}, a^*)] = S[a^*|a_{68}, H(a_{68}, a^*)][h(a^*)]^{D(a^*)}$$

<sup>10</sup> Persons 18 years old and older, including adult children and excluding the spouse if married.

<sup>11</sup> Any children under age 18, regardless of their relationship to the respondent.

<sup>12</sup> Note that the initial age of the first subinterval ( $i = 1$ ) is  $a_0$  and the end age of the last subinterval ( $i = I$ ) is  $a$ .



Because of the proportionality of the hazard function, the denominator probability of surviving to participate in the 1968 survey just cancels the corresponding portion of the numerator probability of observed outcome.<sup>13</sup> Separate models are estimated for men and for women.

## RESULTS

First we present a series of models that allow us to replicate earlier studies on the relationship between marriage and mortality and on sex differences in the married/unmarried differential in mortality. They also allow us to test the hypotheses presented earlier about the effects of financial resources and living arrangements on the mortality advantage of marriage. These models, presented in table 4, include a single dummy variable for current marital status (defined as currently married vs. currently not married, with currently married as the reference category). The first model—the most basic—contains only measures of age and calendar time plus marital status. The second model adds measures of race, own education, education of spouse, region of residence, city size, living arrangements, and household income. The third includes all these except living arrangements, and the fourth all but household income. These latter two models allow us to compare the effects of measures of these two key sets of independent variables with and without the other set in the model.

Model 1 in table 4 allows us to attempt to replicate the result often reported in previous studies of sex differentials in the impact of marriage on mortality. However, we find somewhat different effects than generally reported. Being currently unmarried significantly increases the hazard of dying for both men and women compared to those currently married; the effect is large for both males and females. Although the coefficient for females is somewhat smaller than the coefficient for males (0.52 vs. 0.58), the two are not statistically different from each other. Thus, when we do not hold constant any characteristics of the individuals except their age and the time period of observation, we see no significant difference in the mortality advantage of marriage for men versus women.

Next, in models 2, 3, and 4 we examine the impact of current marital status when we hold constant measures of household income and living arrangements, together with other characteristics of the individual, discussed above. We argued earlier that the married might have lower mortality than the unmarried in part because they have more material resources. The economies of scale and the benefits of specialization in

<sup>13</sup> That is,  $S[a^*|a_{68}, H(a_{68}, a^*)] = S[a^*|a_0, H(a_0, a^*)]/S[a_{68}|a_0, H(a_0, a_{68})]$ .

TABLE 4

## EFFECT OF MARITAL STATUS, HOUSEHOLD INCOME, AND LIVING ARRANGEMENTS ON THE HAZARD OF MORTALITY

VARIABLE	MEN				WOMEN			
	Model 1 <sup>a</sup>	Model 2 <sup>b</sup>	Model 3 <sup>b</sup>	Model 4 <sup>b</sup>	Model 1 <sup>a</sup>	Model 2 <sup>b</sup>	Model 3 <sup>b</sup>	Model 4 <sup>b</sup>
Unmarried . . . . .	.5794*** (.0886)	.6058*** (.1221)	.5999*** (.1212)	.8382*** (.1191)	.5228*** (.1017)	.2835 (.1822)	.2521 (.1802)	.6472*** (.1725)
Household income <sup>c</sup>								
Lowest quartile . . . . .		-.1231** (.0503)	-.1273** (.0500)			-.0952 (.0762)	-.0934 (.0758)	
Second quartile . . . . .		-.3296*** (.1176)	-.3551*** (.1171)			-.6046*** (.1522)	-.6201*** (.1503)	
Third quartile . . . . .		-1.0569*** (.1675)	-1.0784*** (.1662)			-.5357** (.2614)	-.5320* (.2594)	
Fourth quartile . . . . .		-.5492 (.4849)	-.5587 (.4853)			-3.4483** (1.7362)	-3.3940* (1.7359)	
Living arrangements <sup>c</sup>								
Other adults in household . . . . .		-.1855** (.0890)		-.2278** (.0896)		-.1387 (.0883)		-.2262*** (.0867)
Dependent children in household . . . . .		-.0348 (.0474)		-.0901* (.0467)		.0423 (.0569)		-.0194 (.0552)

NOTE.—SEs are in parentheses. Significance is two-tailed from zero.

<sup>a</sup> This model also includes age and calendar time.<sup>b</sup> This model also includes age, calendar time, race, subject's education, spouse's education, region, and city size.\*  $P < .10$ \*\*  $P < .05$ \*\*\*  $P < .01$ .

marriage could mean better health, and this could lower mortality. This reasoning also suggests that women might benefit more from the increased financial resources available to the married than men do, since women's earnings tend to be substantially lower than men's, and the financial well-being of formerly married women is generally lower than that of formerly married men. If this is the case, then we expect the effect of current marital status to decrease when we take household income into account. Models 2 and 3 of table 4, which include a series of splines in (log) household income, allow us to test this reasoning. Recall that model 2 includes all variables, whereas model 3 does not include living arrangements.

When we compare the coefficient for being unmarried in model 4 with the coefficient for being unmarried in model 3, we see that holding financial resources constant reduces the coefficient for marital status for women to insignificance (from 0.65 to 0.25). For men, adding these measures to the model reduces the coefficient from 0.84 to 0.60, but the coefficient for the variable unmarried remains significant. Adding measures of living arrangements to the equation has virtually no impact on the coefficients for household income (compare models 2 and 3). These results suggest that the greater financial resources available in marriage are not responsible for the positive effect of marriage on longevity for men but seem to account for a good deal of it for women.

We also argued that being married may reduce mortality by integrating individuals into a wider network of exchange and support, which may improve the ability to cope with stress and reduce risk taking. We index social integration through coresidence with adults—including a subject's own children—other than the spouse and with any children under age 18.

A comparison of the coefficient for being unmarried in model 1 and in model 4 shows that including measures of the living arrangements of the individual (but not household income) does not reduce but increases the coefficient for being unmarried for both sexes, suggesting that the social integration that comes from living with others (besides the spouse) does not account for the mortality-reducing effect of marriage.<sup>14</sup>

These results point to the sizable differences in the hazard of dying by current marital status. They also show no gender difference in the longevity advantage that married persons show. We see later that, among the unmarried, marital history *does* have different effects on men's and women's chances of dying.

<sup>14</sup> In results not presented here, we find that adding an interaction between current marital status and the presence of other adults and the presence of dependent children has no impact on the coefficient for unmarried for either sex.

*Financial well-being and living arrangements.*—In the section above we discussed how holding financial well-being and living arrangements constant affects the estimated impact of marriage on mortality. Next, we discuss the effect of each of these on mortality directly. As the reader will note, models 2 and 3 show that higher financial resources for the household do significantly lower the hazard of dying for both sexes. The patterns of effects differ somewhat for men and women, however. For men, statistically significant decreases in the risk of dying occur at levels below the top category used here—more than one standard deviation above the mean of household income for all men. For women, increases in the bottom category have little effect, but increases in income within the highest category greatly reduce the risk of dying. Household income appears to be an important predictor of mortality for both women and men, but the largest effects occur at different points in men's and women's income distribution. Comparing models 2 and 3 shows that the effects of household income on the hazard of mortality do not operate through living arrangements; the effects are virtually identical with and without living arrangements included in the model.

Next we turn to the impact of living arrangements on the hazard of mortality for men and women. We begin with model 4, which does not include measures of household income. The results for model 4 show that, for both men and women, having another adult living in the household significantly reduces the hazard of dying. The coefficient for another adult is similar for men and for women. Recall that this person could be a parent, a sibling, an adult child, another relative, or an unrelated individual. The number of dependent children also significantly decreases the hazard of mortality but only for men. When we hold household income constant (model 2), we find that the presence of either adults or children has no significant effects for women. For men we find that only the presence of other adults remains significant. These results suggest that living arrangements have a modest effect on the risk of dying, but that—especially for women—some part of this effect is due to the greater financial resources of larger households.

In results not shown here we tested the hypothesis that the effect of living with others besides the spouse was larger for the unmarried than for the married, since the married, by definition, have at least one other person living with them. Thus, the marginal effect of another adult may be larger for the unmarried. To model 2 we added interactions between being unmarried and the two measures of living arrangements. These results showed no significant interaction between living with another adult and current marital status. In fact, when we included these interactions in the model, we found no statistically significant effect of living arrangements for any group, with the single exception that for currently

unmarried women the effect of living with dependent children on the risk of dying is significantly greater than the (statistically insignificant) effect for married women. This effect appears net of the impact of the generally lower household incomes of female-headed families. Perhaps the stress of heading a household alone increases the chances that a mother dies, or selection of those mothers in the worst health into single parenthood could be responsible for this result. In either case, these findings suggest that cohabitation with another adult does not appear to protect men or women from the increased mortality risk that accompanies singleness.

To summarize briefly to this point, we find a significant and sizable mortality disadvantage for both men and women who are not married compared to the married. But unlike much of the previous literature on the topic, we do not find a significantly larger marriage advantage for men than for women. Our results also suggest that the increased financial resources often accompanying marriage account for much of the marriage effect for women but not for men. For neither sex do living arrangements—either living with other adults or with dependent children—seem to account for any of the lower mortality of the married.

*Marital duration and previous marriage.*—Next, in table 5 we decompose the currently unmarried into the never married, the separated, the divorced, and the widowed. Model 1 contains a dummy variable for unmarried (relative to currently married), together with a series of dummy variables for separated, divorced, and widowed (relative to the never married). These models also include measures of age, calendar time, race, own education, spouse's education, region, city size, living arrangements, and household income. To the extent that those in the worst health are unlikely to marry at all, we expect the never married to be disadvantaged relative to both the currently married and the previously married. In addition, the previously married may retain, at least for a time, some of the protective effects of better health that marriage bestows. If this is the case, the mortality risk of the previously married may lie between those of the currently married and the never married. However, two processes could act to increase mortality risks of the previously married. First, the emotional trauma of the end of a marriage may cause poorer health habits, at least in the short run, and may increase risky behavior; Rosenbloom and Whittington (1993) find that widowhood produces negative effects on eating behaviors and nutrient intakes, for example. Second, married persons with poor health habits and risk-taking behaviors may be selected into the previously married state, either through marital disruption or death of their spouse.

Table 5 shows strikingly different patterns of mortality hazards between men and women for those not currently married. Previously mar-

ried men face mortality prospects that are the same as if they had never been married. The widowed, the divorced, the separated, and the never married all face approximately equal risks of dying and these risks are significantly higher than those faced by currently married men. Currently married women still face the lowest risks of dying, with those who never married and those who are divorced facing significantly higher and approximately equal mortality risks. However, separated women are significantly more likely than the never married to die, and widowed women are significantly less likely. In fact, the mortality risks of currently married and widowed women appear to be quite similar. These results show the uniquely disadvantaged position of separated women. They also point to the favored position of women whose husbands have died relative to other unmarried women. Recall that these effects of marital status hold constant household income, which suggests that, although widowed women often suffer a fall in income with their husband's death (Holden et al. 1988), if their income position remained the same their risk of dying would not go up. This is not the case for never married, divorced, or separated women.

In model 2 of table 5 we test the hypothesis that the protective effects of marriage cumulate over time, as individuals reap the benefits of a more settled, less risky, and therefore healthier lifestyle. This model includes an effect of being currently married versus never married versus previously married, as did model 1, and includes a measure of marital duration. For men, marriage results in an immediate and substantial reduction in the risk of dying—which could result from dramatic reductions at marriage in the prevalence of risky behaviors or from the selection of the healthiest individuals into marriage. For women we see no change in the risk of dying upon marriage, which is consistent with sex differences in the prevalence of risky behavior among the unmarried. However, for both sexes, the hazard of dying falls significantly with marital duration, suggesting a cumulation of the benefits of marriage over time. These benefits could result from the financial advantages of marriage and the better health or health care it brings, from better nutrition and a more settled lifestyle, or from the social and psychological support that comes with marriage. Note, however, that the decrease in hazard of dying that comes with marital duration is significantly larger for women than for men. We speculate later about the causes of these effects.

Although this article focuses on the relationship between marital status and mortality, we briefly discuss the effects of the other predictor variables on the probability of dying. This discussion relies on results from our most complete model, model 2 in table 5. As others have found, the risk of mortality increases with each year of age from age 10, the youngest

TABLE 5

EFFECT OF MARITAL STATUS, MARRIAGE DURATION, HOUSEHOLD INCOME,  
AND LIVING ARRANGEMENTS ON THE HAZARD OF MORTALITY

VARIABLE	MEN		WOMEN	
	Model 1	Model 2	Model 1	Model 2
UnMar(a):				
Unmarried (vs married, zero duration) . . . . .	6578*** ( 1418)	3930** ( 1693)	4559** (.1933)	— 0432 (.2458)
Separated (vs. never married) . .	.0712 ( 2891)	.0818 (.2862)	1.1395*** ( 2856)	1.2124*** (.2802)
Divorced (vs never married) ....	— 2591 (.2448)	— .2417 (.2420)	1192 (.3359)	2165 (.3355)
Widowed (vs never married) ..	— .1014 (.1677)	— 1492 ( 1667)	— .7308*** ( 1568)	— .7542*** (.1567)
MarDur(a) (marriage duration, zero if unmarried) . . . . .				
		— .0085*** ( 0030)		— 0148*** (.0044)
Age(a):				
Intercept . . . . .	— 6 6957*** (.7153)	— 6.5192*** ( 7147)	— 9.6973*** (1.2744)	— 9 5552*** (1.2762)
10–40 years old . . . . .	.0944*** ( 0166)	.0927*** ( 0167)	.1601*** (.0302)	.1613*** ( 0302)
40–60 years old . . . . .	.0798*** (.0114)	.0853*** (.0115)	.0683*** (.0158)	.0780*** (.0161)
60–75 years old .. . . .	.0313*** ( 0097)	.0354*** (.0097)	.0550*** ( 0121)	.0594*** (.0122)
75 or more years old . . . . .	.0590*** (.0123)	.0548*** ( 0123)	.0622*** (.0128)	.0606*** (.0127)
Time(a) . . . . .				
	.0046 ( 0096)	.0085 (.0097)	.0128 (.0119)	.0172 (.0119)
Household income:				
Lowest quartile . . . . .	— 1215** ( 0504)	— 1206** ( 0502)	— .0702 (.0760)	— .0651 ( 0759)
Second quartile . . . . .	— 3292*** (.1177)	— .3116*** ( 1176)	— .6082*** (.1545)	— .5765*** (.1553)
Third quartile . . . . .	— 1 0557*** ( 1679)	— 1 0720*** (.1686)	— .5225** (.2630)	— .5610** (.2642)
Fourth quartile . . . . .	— .5559 ( 4872)	— 5509 ( 4867)	— 3.4461** ( 1 7370)	— 3.4799** (1.7378)
Living arrangements:				
Other adults in household .....	— 1881** (.0892)	— 1868** ( 0892)	— 1459* (.0881)	— 1393 (.0881)
Dependent children in household .	— .0360 ( 0479)	— .0438 (.0484)	.0473 ( 0571)	.0452 (.0572)
Demographics				
Black . . . . .	1416 ( 1305)	1165 (.1316)	.1173 (.1617)	1204 (.1618)
Education < 12 . . . . .	.0988 ( 1009)	.0979 (.1011)	3293*** (.1247)	.3217** ( 1250)
Education ≥ 16 . . . . .	.2309* ( 1253)	.2353* (.1255)	.0893 ( 1572)	.0665 (.1583)

TABLE 5 (Continued)

VARIABLE	MEN		WOMEN	
	Model 1	Model 2	Model 1	Model 2
Spouse's education < 12 <sup>a</sup> . . . . .	.4891*** (.1115)	4893*** (.1131)	1119 (1846)	.1514 (.1842)
Spouse's education ≥ 16 <sup>a</sup> . . . . .	— 1479 (.1700)	— .1829 (.1704)	2254 (2348)	2512 (2354)
Northeast region . . . . .	— .1877* (.1017)	— .1694* (.1016)	.0592 (.1157)	.0721 (.1150)
Large city . . . . .	.3654*** (.0955)	.3701*** (.0949)	.0022 (.1167)	.0118 (.1173)
Very large city . . . . .	.3556*** (.1020)	.3562*** (.1020)	.0001 (.1271)	.0061 (.1277)
Log likelihood . . . . .	—4,271.6	—4,267.1	—3,077.2	—3,071.8

NOTE — Variables are defined in table 3. SEs are in parentheses, two-tailed test for a significant difference from zero.

<sup>a</sup> If married

\*  $P < .10$ .

\*\*  $P < .05$ .

\*\*\*  $P < .01$ .

age we include here. For females, the increase in the baseline hazard is relatively large for each additional year of age from age 10 to age 40—pointing to the very low risk of death at age 10—with slower increases in the hazard of mortality after age 40. For men we see approximately equal increases in mortality risk with increasing age from a low at age ten through age 60. From age 60 to 75 and from 75 on, the increases in risks with each year of age are significant but modest. Note that the coefficients for increasing age after 75 are quite similar for men and women. In fact, differences between the coefficients of the splines for various ages after 40 are small.

The results in table 5 show that black men and women face risks of dying approximately equal to those faced by their white counterparts, all else equal. This lack of an effect of race on mortality matches results by Zick and Smith (1991). However, we find a significant positive effect of race on both male and female mortality in additional results not shown here that do not include measures of household income. This result suggests that much of the often-observed differential in mortality between blacks and whites arises from differences in their income (see also Mott and Haurin 1985; Rogers 1992).

The effects of education are not symmetric for men and women. For women their own education is important and that of their spouse is not. Women with less than a high school degree have significantly higher mortality. For men the education of their spouses is more important than their own. Men with less educated spouses, less than high school,



experience higher mortality rates. Conditional on their spouse's education, men with higher education are at greater risk. We do not have a good explanation for this anomalous result.

An interesting sex difference appears, however, in the effect of the spouse's and subject's own education. For women, having either a relatively well-educated or a relatively poorly educated husband does not affect their own chances of dying. Men, on the other hand, appear to suffer higher mortality risks if their spouses did not finish high school but to receive no extra benefits if their spouse finished college. And a husband's own education has no effect on his risks of mortality, whereas women's own education, if low, significantly increases their hazard of dying. Perhaps husbands with relatively low levels of education increase their wives' mortality risks primarily through their lower contribution to household income. If the contribution of wives appears mostly in household management and the benefits a well-run home confers and in management of health care for the family, then poorly educated wives may provide less of these benefits to their husbands—and to themselves—than do wives with more schooling. Grossbard-Shechtman (1993) suggests that men with traits that are relatively undesirable relative to their wives' characteristics have to compensate their wives by giving them a larger share of the husband's income. This extra access to material resources could be used by the wife to improve her health and well-being.

Men—but not women—living in large cities face heightened risks of dying compared to those in smaller cities or rural areas. And—for men only—those living in the Northeast appear to be less likely to die than those in other regions.

## CONCLUSIONS

Let us summarize by giving our evidence on the mechanisms through which marriage is associated with lower mortality. First, we find—as have others—that when we compare men and women who are currently married with those who are not, the unmarried face substantially higher risks of dying than the married. We also find—contrary to some other findings—no statistically significant difference between the marriage advantage for men and women. Our results show that although increased income does reduce mortality, it does not appear to be the mechanism through which marriage lowers mortality for men but seems to account for much of the protective effect of marriage for women.

Our test of the hypothesis that marriage protects its incumbents by integrating them socially receives no support; although living with other adults (in addition to the spouse for the married) reduces the chances of dying, controlling for living arrangements has no effect on the coefficient

for unmarried for either sex. When we do not hold constant the income of the household, we find that each adult living in the household reduces the hazard of mortality for both men and women and that dependent children reduce mortality risks only for men. When we take into account the greater financial resources that often accompany extra adults, however, we find that the impact of children disappears for men and that the effect of other adults is reduced. These findings suggest that, as Angel and Tienda (1982) argue, poor families may add adults to the household for the extra income they bring.

In fact, when we hold household income constant we find no effect of living arrangements for women, except that unmarried women living with dependent children show higher chances of dying than comparable married women. By contrast, holding constant the composition of the household makes virtually no difference in the size of the effect of household income on the risk of dying. These results imply that for both sexes, but especially for women, access to financial resources is a key pathway through which marriage improves their well-being and life chances.

Our results show that men benefit immediately from reduced mortality risks at marriage. In addition, their risk of dying drops with marital duration as the marriage progresses. We argued earlier that an immediate fall in the hazard of mortality with marriage could result from a drop in risky behaviors but could probably not result from improved health. The benefits of improved health should appear gradually, over time. However, the selection of those with the best physical and emotional health into marriage could also appear as an immediate reduction in the hazard of mortality, although as Goldman (1993) points out, this need not be the case.

For women we see no decline in the hazard of mortality with marriage, but we do see a sizable fall in this hazard with each year of marriage. If, as our results suggest, women benefit to an important extent from the access to higher household income that marriage gives them, then this income may be buying them access to better health care, better nutrition, better housing, a safer job, less physical and mental stress, and so on. Whatever their source, the advantages of marriage for women tend to cumulate with the length of the marriage and are quite sizable.

We also see that movements out of marriage have different implications for the mortality risks of men and women. For men, the end of a marriage returns them to the level of risk that they faced before they had ever married; they fare no better than they did then, but they also fare no worse, and this is the case whether they divorce or their wife dies. For women, mortality risks are high for those who have never married and, by implication, over the lifetimes of those who remain single. The risk of dying for women remains unchanged at marriage but falls substantially over the course of the marriage. For divorced women the chances

of dying rise with the end of marriage to the same level as for the never married; for separated women they rise to significantly higher levels. However, widowed women fare significantly better than all other unmarried women and about as well as the currently married, at least once we hold constant income and living arrangements.

If, as our results suggest, marriage improves the life chances of women primarily through improving their financial well-being, why do never-married, separated, and divorced women fare substantially worse than widowed women, even once we take household income into account? We speculate that widowed women with the same level of household income as divorced women are actually better off financially, since they more often have access to assets that remain from their marriage, especially a house. Divorced women often lose their house and generally must divide assets with their ex-husband, but widowed women more often retain the family home and other financial assets intact. If so, the superior asset position of widows compared to divorced women or never-married women could account for the mortality differentials we observe.<sup>15</sup> The impact of assets, as well as income, on the mortality of previously married women deserves additional attention.

Our results also show that the effect of marriage depends on its longevity and that the relationship between marital duration and hazard of mortality is quite different for men and women. For men we observe a large initial drop in the risk of dying after the wedding, followed by an additional, gradually accumulating benefit of marriage duration. For women we see no immediate improvement in mortality risks with marriage but a somewhat larger benefit over time from being in the marriage.

These findings give us a good deal of evidence on the processes through which marriage protects its incumbents and on the ways that the impact of marriage differs for men and women. "His" marriage seems to consist of a settled life, improved perhaps by the household management skills and labors of his wife, especially if she has at least a high school diploma. Financial resources improve men's life chances but do not alter the impact of marriage itself on mortality. Part of this picture of "his" marriage comes from the importance to the husband's life chances of his wife's educational attainment.

"Her" marriage seems to offer primarily the benefits of improved financial well-being, with the social integration that is unique to marriage and motherhood having little additional effect. Although we have not directly examined the impact of spells of poverty, our results are consistent with Zick and Smith's (1991) finding of substantial positive effects

<sup>15</sup> We are unable to test this reasoning directly, because the PSID includes detailed information on assets of the type we would require only in some years of the survey.

of such spells on the risk of dying, with marginal effects of poverty twice as large for women as for men. Note that, in contrast to "his" marriage, a married woman's life chances do not improve with the education of her husband if we take household income into account, suggesting that financial well-being is the chief route through which highly educated husbands improve their wives' mortality prospects and poorly educated husbands diminish them.

We need to point out, however, that our conclusions are limited by several features of the analysis. First, although we have followed a large and representative sample of individuals into and out of marriages over a lengthy time period, we have not specifically taken into account the process through which some individuals select or are selected into marriage and others are not. Second, we have assumed that individuals are homogeneous in their tendency to dissolve a marriage and in their prospects for losing a spouse to death, given the covariates included in the analysis. In fact, people who lead unsettled lives while married—drinking, smoking, staying out late—are probably precisely the people who are more likely to find themselves unmarried later. And those people whose married life offers the fewest benefits to them may be more likely to end a marriage and more likely to become widowed. Our future work will address both selectivity into marriage and heterogeneity among those who are married. In spite of these limitations, this research provides evidence in support of the argument that marriage protects its incumbents and offers insights into the mechanisms through which this protection is provided.

## REFERENCES

- Angel, Ronald, and Marta Tienda. 1982. "Determinants of Extended Household Structure: Cultural Pattern or Economic Need?" *American Journal of Sociology* 87 (6): 1360–83.
- Antonucci, Toni C., and Hiroko Akiyama. 1987. "An Examination of Sex Differences in Social Support among Older Men and Women." *Sex Roles* 17 (11/12): 737–49.
- Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.
- Booth, Alan, and David Johnson. 1988. "Premarital Cohabitation and Marital Success." *Journal of Family Issues* 9:255–72.
- Brody, Elaine M. 1985. "Women in the Middle and Family Help to Older People." *Gerontologist* 21:470–80.
- Carter, H., and Paul Glick. 1976. *Marriage and Divorce: A Social and Economic Study*. Cambridge, Mass.: Harvard University Press.
- Di Leonardo, Micaela. 1987. "The Female World of Cards and Holidays: Women, Families, and the Work of Kinship." *Signs* 12 (3): 440–53.
- Goldman, Noreen. 1993. "Marriage Selection and Mortality Patterns: Inferences and Fallacies." *Demography* 30 (2): 189–208.
- Goldscheider, Frances K., and Linda J. Waite. 1991. *New Families, No Families?*

- The Transformation of the American Home*. Berkeley and Los Angeles: University of California Press.
- Gove, Walter R. 1973. "Sex, Marital Status, and Mortality." *American Journal of Sociology* 70 (1): 45-67.
- Grossbard-Shechtman, Shoshana. 1993. *On the Economics of Marriage: A Theory of Marriage, Labor and Divorce*. Boulder, Colo.: Westview.
- Helsing, Knud J., Moyses Szklo, and George W. Comstock. 1981. "Factors Associated with Mortality after Widowhood." *American Journal of Public Health* 71 (8): 802-9.
- Hoffman, Saul D., and Greg J. Duncan. 1988. "What Are the Economic Consequences of Divorce?" *Demography* 25 (4): 641-45.
- Holden, Karen C., Richard V. Burkhauser, and Daniel J. Feaster. 1988. "The Timing of Falls into Poverty after Retirement and Widowhood." *Demography* 25 (3): 405-14.
- Hu, Yuanreng, and Noreen Goldman. 1990. "Mortality Differentials by Marital Status. An International Comparison." *Demography* 27 (2): 233-50.
- Kobrin, Frances E., and Gerry E. Hendershot. 1977. "Do Family Ties Reduce Mortality? Evidence from the United States, 1966-1968." *Journal of Marriage and the Family* 39 (4): 737-45.
- Kotler, Pamela, and Deborah Lee Wingard. 1989. "The Effect of Occupational, Marital and Parental Roles on Mortality. The Alameda County Study." *American Journal of Public Health* 79 (5): 607-11.
- Litwak, Eugene, and Peter Messeri. 1989. "Organizational Theory, Social Supports, and Mortality Rates: A Theoretical Convergence." *American Sociological Review* 54 (1): 49-66.
- Mott, Frank L., and R. Jean Haurin. 1985. "Factors Affecting Mortality in the Years Surrounding Retirement." Pp. 31-56 in *Retirement among American Men*, by H. Parnes, J. Crowley, R. J. Haurin, L. Less, W. Morgan, F. Mott, and G. Nestal. Lexington, Mass.: Lexington Books.
- National Center for Health Statistics. 1988. "Health Promotion and Disease Prevention, United States, 1985." *Vital and Health Statistics*, DHHS Publication no. (PHS) 88-1592. Data from the National Health Survey, no. 169. Washington, D.C.: Government Printing Office.
- Pearlin, Leonard, and Joyce Johnson. 1977. "Marital Status, Life Strains, and Depression." *American Sociological Review* 42 (5): 704-15.
- Retherford, Robert D. 1975. *The Changing Sex Differential in Mortality*. Westport, Conn.: Greenwood.
- Rogers, Richard G. 1992. "Living and Dying in the U.S.A.: Sociodemographic Determinants of Death among Blacks and Whites." *Demography* 29 (2): 287-304.
- Rosenbloom, Christine A., and Frank J. Whittington. 1993. "The Effects of Bereavement on Eating Behaviors and Nutrient Intakes in Elderly Widowed Persons." *Journal of Gerontology: Social Sciences* 48 (4): S223-S229.
- Rosengren, Annika, Hans Wedel, and Lars Wilhelmsen. 1989. "Marital Status and Mortality in Middle-Aged Swedish Men." *American Journal of Epidemiology* 129: 54-64.
- Trovato, Frank, and Gloria Lauris. 1989. "Marital Status and Mortality in Canada, 1951-1981." *Journal of Marriage and the Family* 51 (4): 907-22.
- Umberson, Debra. 1987. "Family Status and Health Behaviors: Social Control as a Dimension of Social Integration." *Journal of Health and Social Behavior* 28 (3): 306-19.
- Umberson, Debra, Camille B. Wortman, and Ronald C. Kessler. 1992. "Widowhood and Depression Explaining Long-Term Gender Differences in Vulnerability." *Journal of Health and Social Behavior* 33 (1): 10-24.
- Zick, Cathleen D., and Ken R. Smith. 1991. "Marital Transitions, Poverty, and Gender Differences in Mortality." *Journal of Marriage and the Family* 53 (2): 327-36.

# The Organizational Context of Criminal Sentencing<sup>1</sup>

Jo Dixon  
*New York University*

This article examines sentencing outcomes in 73 counties in Minnesota to appraise three theoretical approaches to sentencing: the formal legal theory of sentencing, which predicts that legal variables are the primary determinants of sentencing, the substantive political theory, which predicts that legal and social status variables determine sentencing, and the organizational maintenance theory, which predicts that legal and processing variables determine sentencing. The findings demonstrate that the effects of legal variables are important determinants of sentencing irrespective of the organizational context whereas the effects of plea are conditioned by the level of bureaucratization in courts. The results also suggest that racial effects on sentencing are curtailed in the context of sentencing guidelines.

## INTRODUCTION

Drawing from a variety of disciplinary perspectives, a substantive body of knowledge about the social organization of criminal sentencing has evolved. However, the empirical research on criminal sentencing has not produced an agreed-upon theory of criminal sentencing processes or a widely accepted set of determinants of sentencing outcomes.

One of the most controversial theoretical debates centers on whether sentencing processes reflect a formal legal, substantive political, or organizational maintenance rationality: the formal legal theory of sentencing predicts that sentencing is primarily determined by legal variables; the substantive political theory predicts that sentencing is determined by

<sup>1</sup> This research was supported by a grant from the National Institute of Justice and a New York University Vladeck Presidential Fellowship. I am indebted to David Greenberg, Christine Harrington, Wolf Heydebrand, Robert Jackson, David Knoke, Mary Lee Luskin, Bernice Pescosolido, Brian Powell, Marjorie Zatz, members of the New York University Institute for Law and Society, and several *AJS* reviewers for helpful comments. Direct all correspondence to Jo Dixon, Sociology Department, New York University, 269 Mercer Street, New York, New York 10003.

legal and social status variables; and the organizational maintenance theory predicts that sentencing is determined by legal and processing variables. A careful review of the empirical findings on sentencing indicates the degree to which this issue is unresolved. Studies that find minorities receiving harsher penalties than nonminorities interpret this finding as evidence that sentencing processes reflect political oppression or substantive political rationality (Garfinkel 1949; Spohn, Gruhl, and Welch 1982; Thomson and Zingraff 1981). Other studies argue that sentencing processes mirror formal legal rationality and that variations in sentencing can be attributed to legal factors. Thus, racial differentials in sentencing occur because of legal attributes correlated with race, for example, offense seriousness, prior record, number of charges, harm inflicted, and weapon use (Clarke and Koch 1977; Chiricos and Waldo 1975). A third set of studies claim support for organizational maintenance rationality in sentencing and argues that variations in sentencing outcomes are due to a combination of legal factors and processing procedures related to sentence reduction, for example, plea disposition (Brereton and Casper 1982; Nardulli 1979; Bernstein, Kelly, and Doyle 1977). Finally, several recent studies observing both direct effects of legal variables and interactions between race and plea maintain that substantive political and organizational maintenance processes operate simultaneously such that legal factors increase sentences to a greater degree for minority populations and pleas are more likely to reduce sentences for nonminority populations (Zatz 1985).

In this article, I integrate the previous sentencing theories into an organizational context perspective on sentencing (Flemming, Nardulli, and Eisenstein 1992; Eisenstein, Flemming, and Nardulli 1988; Hagan, Hewitt, and Alwin 1979; LaFree 1985). This approach suggests that the predominant rationality in the organization of sentencing varies across courts diverging in the organization of judicial and prosecutorial activities. To assess the validity of the organizational context approach, I use ordinary least squares (OLS) and logistic regression techniques to analyze the effects of racial, legal, and processing variables on imprisonment and sentence length in courts varying in the bureaucratization of the technology used to organize the work of prosecutors and judges.

This research strategy has several advantages over previous ones. First, I address a gap in the theoretical literature by integrating the insights of earlier sentencing theories. While these theories disagree on the role of legal, racial, and plea variables, each assumes that one set of variables affects sentencing in *all* courts. Hence, variations in the organization of sentencing among courts are viewed as nonproblematic. Utilizing the insights of more contemporary studies of courts as formal organizations with variations in bureaucratization, I maintain that courts organize the work of judges and prosecutors by employing various tech-

nologies that vary in their degree of bureaucratization (Eisenstein et al. 1988; Hagan et al. 1979; LaFree 1985; Tepperman 1973). Moreover, I suggest that the level of bureaucratization in these technologies shapes sentencing outcomes (Tepperman 1973; Eisenstein et al. 1988; Flemming et al. 1992). Because the organizational context perspective synthesizes and extends the previous theories of sentencing, it offers a more complex and comprehensive examination of sentencing processes.

Second, I address a gap in the previous empirical literature by introducing direct measures of the bureaucratization of court technology into previous quantitative models of sentencing. Studies that introduce urbanization into models use theoretical arguments about bureaucratization to explain the effects of urbanization, but none of these studies introduce direct measures of bureaucratization into sentencing models with formal legal, case processing, and social status variables. Since the inconsistencies in the findings on the significance of legal, racial, and plea variables in nonbureaucratic and bureaucratic contexts form the basis for current debates, this research provides a format for explaining the research findings now considered to be inconsistent. If the bureaucratization of the technology used to organize the work of judges and prosecutors in various courts affects the organization of sentencing, then its omission in previous models makes the previous inconsistencies in findings understandable. Third, I examine several dimensions of the bureaucratization of judicial and prosecutorial technology that add leverage to a theory that posits the organization of court technology as an important determinant of sentencing processes. Fourth, I examine two dimensions of sentencing decisions, imprisonment and sentence length, separately. Because much of the previous research analyzes only one of these dimensions or fails to differentiate between the two, support or refutation of competing theories is often based on comparing different aspects of the sentencing decision. Since the effects of legal, extralegal, and processing variables may be different for these two aspects of sentencing, it is important to model them separately. Finally, analyses of court variations in sentencing processes, such as those in this paper, not only address previous theoretical and empirical debates. In addition, they provide a framework for developing future theoretical and statistical models that capture the variability and complexity of sentencing processes across courts varying in the bureaucratization of technologies employed to organize the work of court actors.

### THREE THEORETICAL PERSPECTIVES

Using the current perspectives on sentencing, I appraise three theories that purport to explain the determinants of sentencing outcomes. Because



most theories of sentencing are general, they make the assumption that there is a unitary sentencing system in all courts. Hence, they describe the rationality of this unitary system such that individual and case characteristics that determine sentencing outcomes in all courts can be identified.

The first theory of sentencing argues that the organization of sentencing is based on substantive political rationality; the second maintains that the organization of sentencing reflects formal legal rationality; and the third portrays sentencing as a natural system based on organizational maintenance rationality. While not a separate perspective, a fourth approach combines the ideas of the substantive political and the organizational maintenance perspectives and argues that these rationalities converge to dilute formal legal rationality. Taken alone, none of these theories can successfully explain the divergent sentencing systems that form the focus of this paper. However, these general theories have implications for explaining variations in sentencing processes across courts that differ in the degree of bureaucratization in the organization of prosecutorial and judicial activities. After specifying the assumptions of each of the aforementioned approaches and assessing the empirical evidence for each, I use an organizational context perspective to formulate an explanation of the organization of sentencing that incorporates various aspects of the existing theories.

### The Substantive Political Perspective

The first approach views the organization of sentencing from a substantive political perspective. One version of this theory views sentencing as a form of political oppression. Fostered by the classic works of Karl Marx (1859) and Thorsten Sellin (1938) and the more contemporary ideas of Richard Quinney (1970), this version argues that the administration of sentencing is a politically organized system wherein the powerful use the police power of the state to reinforce their privileged position by reducing their legal liability for illegal behavior. This line of reasoning focuses on the influence that social class or social status characteristics have on sentencing outcomes (Quinney 1970). A second version of the substantive political perspective also focuses on the influence that social class or social status factors have on sentencing, but explains these effects in terms of social welfare (Savelsberg 1992; Nonet and Selznick 1978; Levin 1977). According to this version, welfare rationalities have motivated the introduction of a substantive political sentencing structure that can influence the relationship between social status and sentencing. Thus, sentencing in the modern welfare state often deems the conditions associated with offenders of lower economic and social status

as mitigating or aggravating circumstances for determining sentencing outcomes.

Regardless of which version of the substantive political theory one accepts, it is assumed that economic and social, as well as legal, factors influence sentencing. The substantive political theory of sentencing would predict that extralegal factors such as class and race, as well as legal factors, play a role in sentencing. Some of the early empirical research on sentencing finds that minorities such as blacks receive harsher sentences than whites, especially blacks who murder whites (Martin 1934; Johnson 1957; Garfinkel 1949). Although these findings are often interpreted as evidence for the substantive political model of sentencing, their use of bivariate statistics renders them weak at best. Only a few contemporary studies find a direct link between social class and/or race once legal attributes are controlled (Spohn et al. 1982; Unnever, Frazier, and Henretta 1980). Most of the results from recent studies that include class or racial and legal variables find support for the formal legal perspective.

### The Formal Legal Perspective

According to the formal legal approach, the organization of bureaucratic and legal decision making is perceived as a technically rational machine. Accepting Weberian descriptions of modern organization as technically rational and applying Weber's argument to the organization of sentencing decisions, followers of the formal legal perspective submit that formal legal rules govern sentencing decisions via the application of these rules to specific cases. Hence, in addition to being predictable, sentencing outcomes are primarily the result of legal rules and criteria applied equally to all classes and races (Lukács 1971).

From the viewpoint of the formal legal theory, one would expect legal factors to be the major determinants of sentencing outcomes. Thus, offenders with different class or status characteristics committing crimes of equal severity would receive similar sentences.

Unlike earlier bivariate research, most contemporary research employs multivariate models that include social class or race, as well as measures of legal variables correlated with class and race. The results from the majority of these studies indicate that class and race become statistically insignificant once legal attributes such as offense severity and prior record are controlled (Chiricos and Waldo 1975; Bernstein et al. 1977; Burke and Turk 1975). Hence, most contemporary research supports formal legal explanations of sentencing processes. Given the scarce evidence for the substantive political model, Hagan (1974) and Kleck (1981) conclude that offender characteristics such as race contribute little to understand-

ing sentencing differentials. Similarly, in a more recent review of the numerous studies of sentencing, Wilbanks (1987) infers that racial and class variations in sentencing are generally reduced to zero when legal variables are controlled and that claims of a racist criminal justice system in the United States are based largely on myth rather than reality.

While most contemporary models of sentencing include the legal variables needed to test the merits of the substantive political and formal legal perspectives, many of them lack the processing variables needed to simultaneously test the merits of the organizational maintenance theory of sentencing.

### The Organizational Maintenance Perspective

Repudiating the Weberian perception of the organization of sentencing as a technically rational machine, as well as the Marxian perception of it as a political machine, the organizational maintenance perspective depicts the organization of sentencing as a natural system that operates on the basis of "Michels's law." According to Michels (1915), the imperatives of organizational maintenance deflect the system from the pursuit of formal rational goals and result in the development of operating goals by organizational elites. When Michels's law is applied to the sentencing process, the organization of sentencing is perceived to be an organizational maintenance process created by courtroom elites. Because a complex network of ongoing informal relationships among court actors is formed, a cooperative effort to efficiently dispose of cases evolves, with effects not envisioned by the substantive political or formal legal models. Courtroom elites come to share common interests in disposing of cases, and the mutual interdependence that develops institutionalizes the presumption of guilt and plea bargaining (Eisenstein and Jacob 1977).

Moreover, an elaborate incentive system for plea inducement emerges. Numerous scholars elaborate on the various incentives that police, prosecutors, defense lawyers, and judges have for cooperating in the inducement of guilty pleas (Newman 1956; Blumberg 1967; Sudnow 1965). Adopting this approach, the central problem for understanding the organization of sentencing becomes defining the operational rather than the political or formal legal goals of sentencing. Because attention is diverted from political or formal rules, attention is given to the role that informal rules and incentives play in organizational maintenance and survival.

The sentencing theory emerging from this approach defines efficient case disposition as the operational goal that maintains a stable and orderly sentencing system. Because all members of the courtroom work group receive benefits from disposing of cases with minimal effort, incen-

tives emerge that positively reward pleas via reduced sentences. Accordingly, "the state, at the relatively small cost of charge reduction or sentence leniency, gains the numerous advantages of the plea of guilty over a long, costly and always uncertain trial" (Newman 1956). An organizational maintenance approach to sentencing would predict that offenders who plead guilty would be rewarded with shorter sentences than those with trial dispositions. From the viewpoint of organizational maintenance theory, one would expect both legal criteria and plea dispositions to have an effect on sentencing independent of the social status of an offender. Because the emphasis is on organizational maintenance, however, social status characteristics such as class and race would have no significant relationship to sentence outcomes.

#### A Combined Substantive Political/Organizational Maintenance Approach

Claiming that political and organizational maintenance goals are simultaneously operating to displace formal legal rationality, some maintain that politics are institutionalized in organizational practices (Selznick 1966). Hence, bias against minorities in sentencing is an indirect political process that is institutionalized behind the facade of organizational maintenance in courts. According to Chambliss and Seidman (1971), certain categories of offenders, specifically whites, are induced to enter pleas that reduce their sentences, while other categories, specifically nonwhites, are processed in ways that fail to reduce their sentences. Because substantive political and organizational maintenance rationality erode formal legal rationality, processing variables and racial variables interact to affect sentencing. Although legal variables affect sentencing, there is also an interaction between the effects of class or race and plea. Hence, minorities or lower-class individuals with plea dispositions do not receive the sentence reductions given to whites or upper-class individuals who plead guilty.

Findings from research incorporating processing variables such as plea into models with racial and legal variables indicate that plea directly affects sentencing or that race and processing modes interact to affect sentencing (Hagan 1975; Lizotte 1978; LaFree 1980; Bernstein et al. 1977; Spohn et al. 1982; Brereton and Casper 1982; Uhlman and Walker 1980; Zatz 1985). Hence, the results are interpreted as evidence for either the organizational maintenance or the combined substantive political/organizational maintenance theories. While these models contain the legal, racial, and processing variables necessary to determine which of the aforementioned theories of sentencing is supported, they are not capable of answering questions about court variations in sentencing processes.

### The Organizational Context Perspective

Taken alone, none of these sentencing theories can fully explain variations in sentencing across courts. Since each perspective begins with the assumption that courts operate with a unitary system of sentencing, some rationality of sentencing is stressed with the exclusion or diminution of others. An alternative and less-utilized approach to the study of sentencing, the contextual approach, provides an analytical tool for capturing variations in sentencing processes across courts. A contextual perspective on sentencing maintains that the sentencing of individuals in a given court is influenced by the political, social, and organizational context of the court. If we accept this premise, it is possible to synthesize the theoretical ideas from the existing models of sentencing and formulate a contextual perspective that ignores neither the formal legal nor the political and organizational maintenance rationalities operating in sentencing processes (Savelsberg 1992).

Reviews of the sentencing literature indicate the need for a contextual analysis of sentencing processes (Hagan and Bumiller 1983; Farrell and Holmes 1991; Savelsberg 1992). Although the few studies of the context of sentencing primarily focus on the influence of social and political environments outside the court on sentencing (Balbus 1977; Levin 1977; Myers and Talarico 1986), several recent studies employ an organizational framework that explores the effects of the formal and informal organizational contexts of courts on sentencing (Hagan 1977; Hagan et al. 1979; LaFree 1985; Nardulli, Eisenstein, and Flemming 1988; Eisenstein et al. 1988; Flemming et al. 1992; Myers and Talarico 1986, 1987). For example, in an analysis of the sentencing of white-collar offenders, Benson and Walker (1988) suggest that differences in court contexts explain the incongruencies between their findings and the earlier findings of Wheeler, Wiseburd, and Bode (1982). A debate continues in the organizational context literature on criminal sentencing over the role that the bureaucratic organization of court work plays in determining which of the various theories of sentencing is supported in nonbureaucratic and in bureaucratic contexts. Most of the empirical work addressing this problem examines the effects of legal variables, social class or status, and plea bargaining on sentencing in rural and urban courts.

One set of studies finds evidence for the formal legal model of sentencing in urban courts and evidence for the substantive political model in rural courts (Hagan 1977; Pope 1976; Austin 1981; Miethe and Moore 1985; Benson and Walker 1988). In other words, they find that legal variables predict sentencing in urban courts, while legal and social status variables affect sentencing in rural courts. Applying Weberian theories of bureaucratic organization, rural-urban differences are attributed to the rationalization and bureaucratic administration of justice found in

urban courts as opposed to the lack of bureaucratic administration in rural courts. According to Weber (1954), the administration of law has become increasingly bureaucratic and rational with abstract formal legal criteria replacing irrational criteria in decision making. More specifically, "whatever cannot be construed in legal terms is . . . legally irrelevant" (Weber 1954, p. 64; Myers and Talarico 1986). In recent formulations of these ideas, Tepperman (1973) makes Weberian assumptions about the organization of sentencing in urban courts and argues that sentencing decisions in urban bureaucratized courts are characterized by universalistic legal standards and a disregard of statuses ascribed to offenders as criteria upon which decisions are made. Antithetically, sentencing in less-bureaucratized rural courts depends on both legal criteria and legally irrelevant criteria such as social status.

In contrast, findings from other recent studies refute these results. Data from several studies demonstrate the opposite pattern; the formal legal model of sentencing fits sentencing practices in rural courts, while the substantive political or organizational maintenance models are reflected in sentencing processes in urban courts (Myers and Talarico 1986, 1987; LaFree 1985; Nardulli et al. 1988; Eisenstein et al. 1988; Flemming et al. 1992). This research finds that legal variables affect sentencing in rural courts, while legal variables and social status or processing variables affect sentencing in urban courts. Two explanations are given for the patterns found in these studies (Hagan 1977; Pope 1976; Austin 1981). One explanation (Chambliss and Seidman 1971) submits that the administration of law in highly bureaucratized courts, such as organizationally strained urban courts, entails the use of institutionalized class, race, and sex inequality. Hence, individuals belonging to politically powerless groups sentenced in strained and bureaucratic urban courts receive more severe punishments than similar individuals sentenced in nonbureaucratic rural courts. The second explanation, based largely on qualitative data on nine courts, maintains that the coupling and informal relations of work groups in rural and urban courts varies in ways that influence decision making such as sentencing. According to this perspective, the nonbureaucratic organization of tightly coupled judicial and prosecutorial activities in rural courts creates stable local legal cultures with high levels of interaction, visibility, and community accountability (Nardulli et al. 1988; Eisenstein et al. 1988; Flemming et al. 1992). As Eisenstein et al. (1988) state, "These features combine to make the court community members more visible to one another, to clients, and to the community at large, increasing accountability to the community and reducing discretion in court processing" (Eisenstein et al. 1988, p. 269). In contrast, work groups in highly bureaucratic and loosely coupled urban contexts characterized by low interaction, visibility, and accountability are more

likely to produce informal local legal cultures that use discretion in decision making. Hence, they employ plea bargaining and sentencing practices that are more individually determined, more discretionary, and more likely to invoke the use of legally irrelevant criteria in sentencing. The evidence supporting this explanation is less than unequivocal, however, since Hagan et al. (1979) find that variations in the coupling between court offices are largely ceremonial, with little influence on sentencing. Whether one accepts the explanation of Chambliss and Seidman (1971) or the alternative explanation of Eisenstein et al. (1988), both predict that rural nonbureaucratic courts are characterized by universalistic legal standards that reduce the use of nonlegal criteria in sentencing. Concomitantly, sentencing in urban bureaucratized courts is more discretionary and allows for the use of both legal and legally irrelevant criteria.

How can one account for the discrepancies between these sets of findings? While previous studies are noteworthy for considering the organizational context of sentencing, they have not directly specified the role that bureaucratization plays in conditioning the effects of formal legal, social, and processing variables on sentencing in nonbureaucratically and bureaucratically organized courts. In general, the prior quantitative research attributes differences in sentencing in rural and urban courts to bureaucratization, without introducing any direct measures of bureaucratization in their models. For example, Hagan (1977) employs urbanization as a proxy for bureaucratization, while Myers and Talarico (1986, 1987) utilize caseload. As Myers and Talarico (1986) suggest, urbanization and caseload are often correlated with bureaucratization, but do not directly tap the dimensions of bureaucratization, such as complexity of division of labor and decentralization of decision making, identified in the organizational literature on bureaucracy and the administrative literature on court organization. Even the presumed strong correlation between bureaucratic organization and urbanization becomes suspect when one looks at the literature on court administration, which shows that bureaucratic organization may be more characteristic of courts in rural areas. For example, courts in rural areas often share judges and devise a very complex division of labor and a very decentralized decision-making process (Luskin and Luskin 1986).

A few studies actually introduce several measures of bureaucratization related to the division of labor and decentralization of decision making in judicial and prosecutorial work, but the data in these studies are limited to cases processed in nine middle-sized jurisdictions (Nardulli et al. 1988; Eisenstein et al. 1988; Flemming et al. 1992). Moreover, their exclusive reliance on court-level data and bivariate analyses of relationships between court characteristics and aggregate sentencing outcomes does not permit them to disentangle the effects of individual- and court-

level variables on sentencing. The degree to which the effects of individual-level variables such as severity, prior record, plea, and race on sentencing vary across courts differing in the organization of judicial and prosecutorial activities has not been adequately addressed and warrants further investigation.

It is premature to accept either set of previous contradictory findings regarding the role of bureaucracy in structuring sentencing processes. The degree to which bureaucratic organization structures sentencing processes is still problematic. To date, no research has examined the effects of formal legal, social status, and processing variables on sentencing outcomes in courts that vary across different dimensions of bureaucratic organization. Thus, this study extends the newly emerging research agenda on organizational contexts by presenting a model that allows a consideration of the impact of various aspects of bureaucratic organization in producing different sentencing processes. If Tepperman (1973) has properly specified the relationship between bureaucratic organization and sentencing, then sentencing in nonbureaucratically organized courts depends on legal criteria and extralegal criteria such as social status, while in bureaucratized courts, a reliance on social status will be absent and sentencing will be primarily based on legal criteria such as offense severity and prior record. If, on the other hand, bureaucratic organization structures incentive systems for plea bargaining (Chambliss and Seidman 1971) and/or the informal organization of work-group cultures (Eisenstein et al. 1988), sentencing in loosely coupled bureaucratic contexts will invoke the use of legally irrelevant criteria related to plea or race more than sentencing in nonbureaucratic contexts, where plea incentives are less and coupling among court actors is tightly structured.

In the past few years, there have been fewer sentencing studies that empirically examine sentencing processes in courts with differing organizational contexts. Perhaps, this reflects dissatisfaction with the lack of any coherent pattern in the sentencing literature. Some authors suggest that the development of sentencing guidelines has reduced discretion in judicial decision making and lessened the possibility that social status and processing can contaminate the sentencing process. The first legislatively mandated sentencing guidelines were developed in Minnesota in 1980 (a more detailed discussion of the Minnesota sentencing guidelines is provided in the next section), although most states, as well as the federal government, have implemented some form of guidelines for sentencing (Champion 1989). Rejecting the notion that sentencing guidelines have eliminated sentencing disparities, others argue that the effects of legally irrelevant criteria creep back into sentencing once sentencing guidelines are institutionalized or that discretion is moved from the sentencing stage to earlier stages of processing (von Hirsch, Knapp, and Tonry 1987;



Greenberg and Humphries 1980). While it would be premature to generalize the findings in guidelines states to nonguidelines states, the predominance of guidelines in most states compels scholars interested in sentencing to examine sentencing in guidelines states as well as in nonguideline states. While there are numerous studies in nonguidelines states, there are only a few studies in states with sentencing guidelines. Moreover, these studies generally ignore organizational variations among courts that may produce differential sentencing patterns despite the introduction of sentencing guidelines.

To date, no study has been conducted that examines the context of sentencing where an explicit set of sentencing guidelines are implemented across courts varying in the degree of bureaucratization in the organization of prosecutorial and judicial activities. This research attempts to fill this void by extending the newly emerging agenda on contextual effects by looking at organizational contexts in a state with sentencing guidelines. While it is premature to generalize from the findings on sentencing in one state, this research points to the importance of state sentencing policies and organizational variations among courts for understanding the contradictory findings of previous research on sentencing processes. If there are variations in sentencing processes across differentially organized courts in a sentencing guidelines state, then there will certainly be variations in states with no state-organized effort to reduce the influences of social status and processing factors on sentencing.

#### THE DATA

The sample of 1,532 cases for this study comes from the Minnesota State Court Administrator's case-tracking system and consists of all felony cases ending in conviction initiated in the first six months of 1983.<sup>2</sup> The data come from four sources. (1) Information on the offense severity, prior record, race, sentence type, and sentence length comes from the Minnesota Sentencing Commission's database for convicted cases in Minnesota.<sup>3</sup> (2) Attorney-type and disposition-type information comes from the statewide case-tracking system of the Minnesota State Court Administrator's Office. The assistant director and staff members of the Minnesota State Court Administrator's Office and the director and staff mem-

<sup>2</sup> Because the percentage of cases censored, i.e., not completed before the end of the data collection period, is extremely small and the residuals from an OLS regression approximate the normal distribution, the censored cases in this sample do not present problems for estimation.

<sup>3</sup> There is no problem with missing data on the case information variables since the number of cases with missing data is minuscule.

bers of the Minnesota Sentencing Guidelines Commission provided invaluable assistance in preparing the data for analysis and spent numerous hours explaining the Minnesota court rules and sentencing guidelines. (3) The information on the organization of the work of the judges comes from responses to telephone interviews with the judicial administrators of the 10 judicial districts in the state of Minnesota. Interviews with the Minnesota assistant state court administrator informed us of the types of variations in the organization of judicial activities in Minnesota. The Minnesota State Court Administrator's Office provided maps of each of the 10 judicial districts and the names of the judicial administrators for each of the 10 districts. They also contacted the judicial administrators and informed them that they would be receiving a phone call from us and that they were to provide us with information on the types of docket and calendaring systems used by judges in each county. We conducted telephone interviews with all of the judicial administrators. They provided us with the information on docket and calendaring systems in each county in Minnesota, as well as other information on judicial assignments. (4) The information on the organization of the prosecutors' offices comes from responses to a questionnaire given to the chief prosecutor in each county. Before designing the questionnaire, we interviewed the director of the Minnesota County Attorney's Association, a previous court administrator for Minneapolis, and several previous chief prosecutors in Minneapolis to gain information about the various ways the prosecutors organize the work in their offices. In addition, several judicial administrators and prosecutors pretested the prosecutor questionnaire on the organization of prosecutorial activities. The director of the Minnesota County Attorney's Association provided us with the name and address of the chief prosecutor in each county. We mailed a revised questionnaire to the chief prosecutor in each county. In order to better understand the organization of each prosecutor's office, we asked them to return any existing written organizational chart of their office with their questionnaire. For offices with no existing organizational chart, we asked the chief prosecutor to sketch a chart indicating departments, divisions, and lines of authority. After receiving the returned questionnaires, we conducted interviews with the chief prosecutors in Minneapolis and Saint Paul.<sup>4</sup> It was our intention to examine organizational variations in the public defenders' offices, but an interview with the director of the Minne-

<sup>4</sup> Information on the judicial administration was obtained for all of the counties in Minnesota, while information on the prosecutors' offices was obtained for 84% of the counties. A careful examination of other characteristics of the missing courts gives no evidence of selection bias. These courts are located in various geographic areas and do not come from any particular judicial district.

sota Association of Public Defenders revealed that their organizational structures were relatively uniform across courts since their control was state rather than county based. Although Hagan et al.'s (1979) work on the influence of probation officers on sentencing is important in some contexts, it is not applicable in Minnesota since the introduction of sentencing guidelines in Minnesota was accompanied by the elimination of probation officer sentencing recommendations.

It is important to note that Minnesota is a sentencing guidelines state. In fact, Minnesota was the first state to implement both presumptive dispositional and durational guidelines for sentencing. The guidelines went into effect for any felony conviction after May 1, 1980 (for a detailed description of the guidelines, see Miethe and Moore [1985], Martin [1983], Minnesota Sentencing Guidelines Commission [1984], Knapp [1982], and von Hirsch et al. [1987]). The intent of the guidelines was to establish a set of consistent standards for sentencing that would increase uniformity in both the dispositional (in/out) and durational (sentence length) aspects of sentencing decisions. A "modified deserts" standard was adopted whereby presumptive sentencing dispositions and durations are determined by the severity of the conviction offense and, to a lesser extent, the criminal history of the offender. The presumptive sentencing dispositions and durations for each combination of severity/criminal history were determined by the Minnesota Sentencing Guidelines Commission and are presented in the form of a sentencing grid.

Although the Minnesota system is considered one of the most comprehensive determinate sentencing systems, there is sufficient opportunity for sentencing disparities. For example, judges may diverge from presumptive sentences within a 15% range without departing from the guidelines. In one evaluation, Miethe and Moore (1985) find that the likelihood of receiving an aggravated departure from the guidelines is influenced by the socioeconomic attributes of defendants. While evaluations of the role of offense severity, prior record, plea, and race under the Minnesota guidelines have been performed (Minnesota Sentencing Guidelines Commission 1984; Miethe and Moore 1985; Moore and Miethe 1986; Miethe 1987), these previous studies are unable to gauge court variations in the predictors of sentencing outcomes. It is also true that these studies were conducted shortly after the implementation of sentencing guidelines. If, as some authors have suggested, sentencing practices in Minnesota adhered to guideline standards during early implementation, but have shifted toward preguideline patterns in the later years of institutionalization, then it is important to continually monitor sentencing processes (von Hirsch et al. 1987).

Although previous studies provide evidence that the effects of social characteristics such as race and socioeconomic status diminished with the introduction of sentencing guidelines in Minnesota, the extent to which

discretion moved from sentencing to plea bargaining or occurred in courts with particular forms of judicial and prosecutorial organization is unknown. Because sentencing guidelines are designed to minimize discretion in sentencing dispositions and durations across judges and courts, they provide a particularly rich context for studying the variations in the determinants of sentencing across courts. Estimating the effects of social and processing variables on sentencing in a sentencing guidelines context provides the most conservative estimates of their effects. If there are significant differences in the effects of various social and processing variables on sentencing once the effects of legal variables are controlled in guidelines states, then there will certainly be greater effects in states without state-imposed sentencing guidelines. Moreover, if the effects of legal, processing, and social variables vary across organizational settings in a guidelines state, then they will certainly vary across organizational settings in states without state policies designed to create uniformity. While disparities in sentencing are reduced in Minnesota because of the guidelines, organizational variations are maximized. Because counties vary in the organization of judicial and prosecutorial activities in Minnesota, it is possible to compare sentencing processes across courts varying in their organizational forms. In summary, the existence of a set of sentencing guidelines designed to reduce sentencing disparity coupled with county variations in judicial and prosecutorial organization in Minnesota renders it an optimal state for exploring the effects of legal, processing, and social variables on sentencing in courts varying in the bureaucratization of the technology designed to organize the activities of judges and prosecutors.

## THE MEASURES AND METHODS

### Measuring the Dependent Variables

I analyze two measures of the dependent variable, sentencing. One reflects the probability of receiving a prison-term sentence, and the other gauges the length of the sentence received for defendants receiving a prison term. I compute the prison-term sentence by giving a score of "1" to all cases receiving a prison term as all or part of their sentence. This includes cases receiving only prison terms and cases receiving prison terms plus other punishments such as fines and probation. Cases score "0" if they do not receive prison as all or part of their sentence. Included in this category are cases that receive diversion, probation, fines, jail terms, or some combination of these. As for the prison-length measure of sentencing, I measure the number of months of prison given to the defendant at sentencing. The measures for the dependent and independent variables are shown in table 1.

TABLE 1  
CODING PROCEDURES AND SUMMARY MEASURES FOR VARIABLES USED IN  
SENTENCING MODELS

Variable	Coding
Sentencing outcome	
Prison .....	Was the offender sentenced to prison? (0 = no; 1 = yes; $\bar{x}$ = .21, SD = .41)
Sentence length .....	Length of prison sentence in months for offenders sentenced to prison (range = 1-140; $\bar{x}$ = 13.04, SD = 17.98)
Legal:	
Severity .....	Minnesota sentencing guidelines: severity level based on statutes (range = 1-10, $\bar{x}$ = 3.51, SD = 1.78)
History .....	Minnesota sentencing guidelines: total criminal history points based on weighted history (range = 1-6; $\bar{x}$ = 1.30, SD = 1.77)
Multiple charges .....	Were there multiple charges brought against the offender at the original charging stage? (0 = single charge; 1 = multiple charges, $\bar{x}$ = .49, SD = .50)
Weapon .....	Did the offender use a weapon in the commission of his/her crime? (0 = no; 1 = yes; $\bar{x}$ = .05, SD = .22)
Processing:	
Plea .....	Did the offender go to trial or plead guilty? (0 = trial; 1 = plea; $\bar{x}$ = .95, SD = .21)
Attorney type .....	Did the offender use a public or private attorney? (0 = public attorney; 1 = private attorney, $\bar{x}$ = .18, SD = .38)
Social status:	
Race .....	Race of offender (0 = nonwhite; 1 = white; $\bar{x}$ = .79, SD = .41)
Judicial bureaucratization:	
Judicial complexity .....	Was this case processed in a court where the judicial division of labor is characterized as non-complex with mixed dockets or complex with specialized criminal dockets? (0 = noncomplex division of judicial labor [nonspecialized mixed docket]; 1 = complex division of judicial labor [specialized dockets with specific criminal docket]; $\bar{x}$ = .52, SD = .50)
Judicial decentralization .....	Was this case processed in a court where judicial decision making is characterized as centralized with an individual judicial calendar or decentralized with a master judicial calendar? (0 = centralized judicial decision making [individual calendar], 1 = decentralized judicial decision making [master calendar], $\bar{x}$ = .66, SD = .47)

TABLE 1 (Continued)

Variable	Coding
Prosecutorial bureaucratization:	
Prosecutor complexity .....	Was this case prosecuted in a prosecutor's office where the prosecutorial division of labor is non-complex or complex? (0 = noncomplex division of prosecutorial labor [three or fewer divisions or departments]; 1 = complex division of prosecutorial labor [more than three divisions or departments]; $\alpha = .65$ , SD = .48)
Prosecutor decentralization .....	Was the case prosecuted in a prosecutor's office where the prosecutorial decision making is centralized in the hands of the chief prosecutor or decentralized so that any assistant prosecutor can make decisions? (0 = centralized prosecutorial decision making [only chief prosecutor has authority to make plea bargains]; 1 = decentralized prosecutorial decision making [any assistant prosecutor has authority to make plea bargains]; $\alpha = .78$ , SD = .41)

### Measuring the Independent Variables

I use four measures of the legal variables. The first has a range of 1–10 and conveys the severity of the conviction offense. Several first-degree murder cases are excluded because they are outliers, since the sentence is mandatory imprisonment for life. Previous research suggests a number of possible measures of offense severity: maximum sentence allowed by law, judge's evaluations of seriousness, public survey response evaluations of seriousness, number of defendants, whether defendants resist arrest, degree of injury to victims, and whether weapons are used (Hagan, Nagel, and Albonetti 1980; Lizotte 1978; LaFree 1980). The measure I employ indicates the seriousness level of the most severe conviction offense as determined by the Minnesota Sentencing Guidelines Commission (1984).<sup>5</sup> The commission assigns every Minnesota felony statute a severity score with a range of 1–10.<sup>6</sup>

<sup>5</sup> When an offender is convicted of more than one charge, the severity level of the most serious charge is used since this is the charge used for determining the sentence for offenders convicted of multiple offenses.

<sup>6</sup> Because the severity level I use is the same seriousness level used by the courts to sentence, it has less measurement error than measures in previous studies, where less is known about the dimensions used in formulating the seriousness level used by court actors. However, it should be noted that the severity level is organizationally derived by the Minnesota Sentencing Guidelines Commission.

The second legal variable operationalizes the criminal history of the offender. Previous studies employ a number of scientifically derived measures of prior record. Some commonly used ones are prior arrests, prior convictions, and prior incarcerations (Lizotte 1978; Zatz 1985). A few studies include prior juvenile arrests (Greenwood, Petersilia, and Zimring 1980). Rather than a scientifically derived measure, the criminal history measure I use is an organizationally derived measure (Minnesota Sentencing Guidelines Commission 1984). This index, ranging from one to six, is computed by summing the weighted points accrued from the following items: (1) prior felony record, (2) custody status at the time of the offense, (3) prior misdemeanor and gross misdemeanor record, and (4) prior juvenile record for young adult felons (for a detailed discussion of the weighting of criminal history information, see Minnesota Sentencing Guidelines Commission [1984]).<sup>7</sup> Although the sentencing commission specifies the items to be used in computing the prior-record scores, it has been noted that prosecutors often engage in a practice of building up the prior criminal history scores to ensure imprisonment (von Hirsch et al. 1987). In order to avoid biases introduced by this practice, I am using the prior-record score determined by the sentencing commission at disposition rather than by the prosecutor at the charging stage.

The third legal variable, multiple charges, is a dichotomous variable that designates whether the defendant had one charge or multiple charges at the original charging stage. Finally, the fourth legal variable, weapon, is operationalized as a dichotomous variable designating whether the offender used a weapon in the commission of his or her crime.

In appraising the influence of legal variables on sentencing, I expect each to have a positive effect on the probability of imprisonment and the length of sentence. The Minnesota sentencing guidelines explicitly incorporate conviction offense severity and prior history as the basis for the prescriptive sentences recommended to judges. Because Minnesota defines retribution as the primary task of sentencing, I expect the severity of the conviction offense to have more explanatory power than the other legal variables. Although the Minnesota Sentencing Commission failed to adopt utilitarian goals, this goal was retained with reduced importance with the inclusion of prior record as the lesser of the two major components of the recommended sentence. Hence, I expect prior record to affect sentencing, but to a lesser degree than offense severity. Since the use

<sup>7</sup> The measure of criminal history used in this research is more detailed than that used by previous researchers because it weights the previous offenses according to seriousness rather than adding up the number of previous prior offenses with no regard for their seriousness. It also contains less measurement error because it is the score used by the sentencing commission in determining presumptive sentences.

of a weapon and the perpetration of multiple offenses are aggravating circumstances that are legitimate reasons for increasing sentences, it is expected that they will have a positive effect on sentencing, although this effect will be less than that of the two major determining factors, offense severity and prior record.

To operationalize case processing, I use several variables. The first is a plea variable that differentiates cases resolved by pleas from cases resolved by trial by jury or trial by judge. Guilty-plea cases score "1" and trial cases, whether by judge or jury, score "0."<sup>8</sup> The common claim that cases represented by private lawyers follow a different route in case processing prompted the inclusion of attorney type as a second case processing variable. Cases using a private attorney receive a score of "1," and cases using a public attorney or representing themselves receive a score of "0."<sup>9</sup>

As in most of the previous studies, I examine the effects of race on sentencing to assess the role of social status on sentencing and to test the merits of the substantive political theory of sentencing. Measures of either race or income are typically included since there are usually problems of multicollinearity if both measures are introduced in the same model. Since I do not have any available data on the income or occupational status of offenders, I include race in my models. Race is often included rather than income because information on race is generally more available in court data, but also because it is typically a more salient basis for discrimination in criminal court. This is because most criminal offenders are in the lower income range, while there are considerably more variations in race and ethnicity to use as a basis for discrimination. Cases score "0" if they are nonwhite, "1" if they are white.<sup>10</sup> There are several issues regarding the race variable that should be mentioned. First, several studies reveal that the racial composition of the offender/victim dyad is

<sup>8</sup> Although the data contain a substantial number of trial cases, the number is not large enough to distinguish trial-by-jury cases from trial-by-judge cases in statistical analysis.

<sup>9</sup> The number of offenders representing themselves is minuscule. Hence, this variable primarily represents a distinction between the case processing routes of private and public attorneys.

<sup>10</sup> Some sentencing studies compare whites to nonwhites (Lizotte 1978; Chiricos and Waldo 1975), while others conduct separate analyses for whites, blacks, and Hispanics (Zatz 1985). While there are a sufficient number of nonwhites, especially Native Americans, the number in each category is not large enough to distinguish the various nonwhite categories in a statistical analysis, especially in the contextual subsamples. While the size of the race coefficient may vary across the nonwhite categories, the direction of the effects of race should be the same for the minorities. I conducted a separate analysis of the various races in the nonwhite category, and the same patterns of sentencing are found for each group.



an important consideration in sentencing, especially for cases of rape (LaFree 1989; Myers and LaFree 1982). Since I do not have information on the race of the victim, I am unable to introduce this variable into my model. To evaluate the degree to which this exclusion introduces bias, I ran one analysis removing the rape cases and another removing all sex offenses. Because the results from these two analyses are analogous to the findings in this paper, it is unlikely that the missing information on the race of the victim biases estimates of the effects of race. Second, the distribution of race across courts could present potential problems of estimation if all the nonwhite offenders are in several courts. This is particularly important since the research design attempts to assess variations in the effects of variables such as race across courts. An evaluation of the distribution of race across courts reveals that this is not as problematic as might be expected. While it is true that blacks are overrepresented in courts in urban counties such as Ramsey and Hennepin, it is also true that Native Americans are overrepresented in courts in rural counties. Since additional analysis shows that the effects of blacks and Native Americans are similar, I combined them into a nonwhite category. Combining the nonwhite categories reflects the distribution of nonwhites across courts and creates organizational contexts with enough minority offenders to permit reliable estimates. Given that Minnesota is a sentencing guidelines state where minorities constitute a small proportion of the general and criminal population, I expect to find less effects of race than would be found in studies using data from nonguidelines states.

I employ several measures of bureaucratization to test whether the level of bureaucratization of judicial and prosecutorial administration where cases are processed conditions the effects of legal, processing, and social variables on sentencing. The measures are derived from previous research on the organization of judicial and prosecutorial activities (Luskin and Luskin 1986; Eisenstein et al. 1988; Nardulli et al. 1988; Flemming et al. 1992; Hagan et al. 1979; LaFree 1985) and previous research on bureaucratic organization (Hage and Aiken 1969; Blau 1974; Hall, Haas, and Johnson 1967). I, like previous researchers, began with the assumption that urban courts are organized in a bureaucratic fashion while rural courts are less bureaucratized. However, a close examination of the organization of judicial and prosecutorial activities in Minnesota tells a different story. For example, courts in rural areas often display a complex division of labor and decentralization of decision making due to the sharing of judges across counties. This sharing also suggests that describing the local legal culture of rural courts as lowly bureaucratized with tight coupling is suspect. It is often the case that judges hear some cases in counties quite removed from the counties where they reside.

Courts are composed of the set of offices of several key court actors: the judges, the prosecutors, and the defense lawyers. Because the organization of public defenders is relatively uniform across courts in Minnesota, I emphasize variations in the organization of the activities of judges and prosecutors. The organization of judicial and prosecutorial activities varies along a number of administrative dimensions that reflect the components of bureaucratic organization stressed in the literature on formal organizations. As mentioned earlier, this organization may be correlated with urbanization and caseload but in no way bears a one-to-one correspondence. For state courts in Minnesota, each county selects its own judicial and administrative structures.

Despite the common perception that plea bargaining is essential for controlling court dockets and influential in sentencing in courts with heavy caseloads but not operative in courts with small caseloads, empirical studies provide little support for the caseload pressure hypothesis (Heumann 1975; Wooldredge 1989; Heumann and Loftin 1979; Meeker and Pontell 1985; Myers and Talarico 1987; Holmes, Daudistel, and Taggart 1992). Most of these studies find that plea bargaining is high regardless of the caseload of the court (Heumann 1975). The paucity of evidence for the caseload pressure argument, coupled with the problems of using urbanization as a proxy for bureaucratization, spurred researchers to seek other explanations for variations in plea bargaining and sentencing across courts. Alternative explanations focusing on organizational factors have proved fruitful. I introduced a measure of caseload in my models, but it was never statistically significant.<sup>11</sup> Thus, the organizational context remains the focal point of this research.

Variations in the organization of judicial and prosecutorial activities in Minnesota correspond to the components of bureaucratic organization in the literature on formal organizations; variations in the organization of the judicial dockets and the prosecutorial divisions correspond to variations in the division of labor; variations in judicial calendaring systems and prosecutorial plea-bargaining policies correspond to variations in the decentralization of decision making. Treated as variables, these administrative factors can be viewed as indicators of related, but distinct component parts of bureaucratic organization that are present in varying degrees in the technology used to organize the judicial and prosecutorial administrative structures where cases are processed.<sup>12</sup> The conventional

<sup>11</sup> Given that previous empirical research has generated little evidence for the effects of caseload and a great deal of evidence for the effects of organizational factors, I chose to include the organizational factors.

<sup>12</sup> Some of the component parts of bureaucratization are highly correlated with one another, while others are only slightly correlated (see the correlation matrix in the

view that size is an adequate measure of bureaucracy because the various components of bureaucratization covary has been challenged by those who have attempted to develop organizational taxonomies. These studies maintain that bureaucratic organization is not a unitary element, but rather a heterogeneous category with components that need to be specified and measured independently (Haas, Hall, and Johnson 1966; Pugh et al. 1968; Warriner, Hall, and McKelvey 1981). While there is some conceptual clarity regarding the important structural characteristics of bureaucratic organization, the specific referents of these characteristics vary with the type of organization studied. Because the component elements of bureaucratization in courts may have varying effects on organizational processes such as sentencing, I analyze them separately. The measures developed in this paper apply the measures of bureaucratization developed by various organizational researchers and used extensively by Hage and Aiken (1969).<sup>13</sup> Their specific referents are based on previous studies of the organization of courts (Eisenstein et al. 1988; Nardulli et al. 1988; Flemming et al. 1992; Hagan et al. 1979; LaFree 1985; Luskin and Luskin 1986).

Two component parts of the bureaucratic organization of the judiciary are measured: complexity of division of judicial labor and decentralization of judicial decision making. Complexity of the division of judicial labor is operationalized by determining whether the criminal court where a case is processed operates with a noncomplex mixed docket, where all judges hear a variety of cases (e.g., criminal, civil, juvenile), or whether the court operates with a complex specialized docket, where one set of criminal-docket judges hear all the criminal cases. Cases processed in a court with low complexity (mixed docket) in the organization of the docket receive a score of "0," and cases processed in a court with high complexity (specialized docket) in the organization of the docket receive a score of "1."<sup>14</sup>

The second component of judicial bureaucratization, decentralization

---

appendix). Because the literature on organizations suggests that these are related but distinct components of bureaucratic organization, I analyze each separately. Since there are separate equations for each component part and none of the component parts are highly correlated with the independent variables, correlations between the component parts do not present estimation problems.

<sup>13</sup> For a full discussion of these measures, see Hall et al. (1967).

<sup>14</sup> These administrative variables provide a more precise measure of bureaucratic organization than the city-size or caseload-size variables used in previous research. Although city size and caseload are correlated with bureaucratic organization, they fail to capture the variations of bureaucratic organization across cities and across courts with high caseloads. For example, judicial and prosecutorial activities are much more bureaucratically organized in Minneapolis than in Saint Paul, both urban courts with high caseloads.

of judicial decision making, is operationalized by examining the case assignment or calendaring system used in the court where the case is processed. Cases score "1" if they are processed in a court using a decentralized master calendaring system, whereby judges are responsible for the performance of a particular task (e.g., arraignment, bail, plea, or trial) for a case; cases score "0" if they are processed in a court with a centralized individual calendaring system, where judges are assigned cases rather than tasks and are responsible for all the tasks involving a case from beginning to end.

Similarly, two components of the bureaucratization of the prosecutors' offices are measured, prosecutorial complexity and decentralization. To operationalize the complexity of the division of labor in the prosecutor's office where a case is processed, I use a dichotomous variable that gives a score of "0" to cases processed in offices with three or fewer divisions or departments and "1" to cases processed in offices with more than three divisions or departments.<sup>15</sup>

The second component of prosecutorial bureaucratization comes from the responses of prosecutors to a questionnaire about the operation of their offices and indicates whether a case is processed in a prosecutor's office where decision making regarding plea bargaining is centralized or decentralized. Each prosecutor's office designated whether any assistant prosecutor has the authority to make a plea bargain or whether that authority rests solely with the chief prosecutor in the office. By authority, I mean that action can be taken without waiting for confirmation from above, even if, in fact, the decision is later ratified at a higher level. For each case, a "0" score is given if plea bargaining in the prosecutor's office is centralized in the hands of the chief prosecutor, and a "1" score if it is decentralized such that any assistant prosecutor can make a decision about a plea bargain.

If the organization of sentencing is uniform in all courts, I expect it to be characterized in one of the following four ways: (1) if the formal legal theory of sentencing best describes the sentencing process, sentence type and length in all organizational contexts will be influenced largely by legal variables; (2) if in addition to legal variables race is influential while processing variables are not in all legal contexts, there is support for the substantive political theory; (3) if the legal variables and processing variables are influential while race is not in all organizational contexts, support for the organizational maintenance model is given; (4) if legal variables are influential with race interacting with plea to influence sen-

<sup>15</sup> The decision to use three divisions or departments as the cut point for low and high complexity comes from theoretical considerations and an examination of the distribution of the number of departments across courts in Minnesota.

tencing outcomes in all organizational contexts, there is evidence for the combined substantive political/organizational maintenance theory.

On the other hand, if the organizational context perspective is more applicable, I expect that sentencing processes will vary with the organizational context. If Tepperman (1973) is correct, then sentencing in bureaucratic courts will be determined by legal factors, whereas sentencing in nonbureaucratic courts will be determined by legal factors and legally irrelevant factors such as race and processing. If, however, alternative theories are more credible (Chambliss and Seidman 1971; Eisenstein et al. 1988), legal factors will predict sentencing in nonbureaucratized courts, and legally relevant and legally irrelevant factors will predict sentencing in bureaucratized courts.

I use logistic regression to estimate an equation for prison and OLS regression to estimate an equation for sentence length without any information on the organizational context in which cases are processed.<sup>16</sup> Next, I regress the two dependent measures on the legal, processing, and social status measures for cases in each of the judicial and prosecutorial contexts to determine whether the factors that predict prison and sentence length are the same regardless of the organizational context in which a case is processed.<sup>17</sup> Differences between the coefficients for each variable in the low and high categories of each of the judicial and prosecutorial contexts are tested.<sup>18</sup>

<sup>16</sup> The calculation of OLS is not adequate for testing the prison model because the dependent variable is dichotomous. Logit was chosen because it gives consistent estimates for models with dichotomous dependent variables.

<sup>17</sup> The substantive implications of bias due to sample selectivity are of particular concern in estimating the OLS coefficients for sentence length since only 300 of the 1,446 cases sentenced received a prison term. In an effort to determine whether bias was introduced, I estimated two equations: one to model the selection process and the other to model the effects of selection and the other exogenous variables on the sentence length. Following Heckman (1979), probit was used to estimate the first model and OLS regression for the second. Since the results from this analysis shows no significant change in the effects of the other variables on the dependent variable, the interpretations of the results are the same for the uncorrected and the corrected models.

<sup>18</sup> The significance of the difference between any pair of logit or regression coefficients across equations is tested with the following test statistic, which has the following *t*-distribution.

$$t = \frac{b_1 - b_2}{[V_1 (SEb_1^2) + V_2 (SEb_2^2)/(V_1 + V_2)]^{1/2}}$$

where

$b_i$  = logit coefficient for population  $i$ ,

$V_i = N - K - 1$  for population  $i$ ,

$SEb_i$  = standard error of logit coefficient for population  $i$ .

TABLE 2  
UNSTANDARDIZED LOGISTIC REGRESSION ESTIMATES  
FOR PRISON MODEL OMITTING THE ORGANIZATIONAL  
CONTEXT

Variable	Prison Model
<b>Legal.</b>	
Severity . . . . .	.71* (.07)
History . . . . .	1.21* (.07)
Weapon . . . . .	1.93* (.38)
Multiple charges . . . . .	.10 (.20)
<b>Processing.</b>	
Plea . . . . .	-1.37* (.44)
Attorney type . . . . .	-.52 (.30)
<b>Social status<sup>a</sup></b>	
Race . . . . .	-.37 (.23)
<i>N</i> . . . . .	1,446
-2 log <i>L</i> . . . . .	810 *
<i>df</i> . . . . .	7

NOTE — Values in parentheses are SEs.

\* $P < .05$ .

## RESULTS

First, I ask which of the four main theories of sentencing predicts prison and sentence length when the organizational context in which a case is processed is omitted.<sup>19</sup>

The answer given by the results from the analysis predicting the prison decision, shown in table 2, and by the results from the analysis predicting the sentence length decision, shown in table 3, is the organizational maintenance theory. When I regress prison on racial, legal, and processing variables, only the legal variables and the plea variable are statistically significant predictors of prison. Increases in the severity of the disposition charge and criminal history, as well as the possession of a weapon, sig-

<sup>19</sup> The interactions between plea and race were introduced in the various models to test the combined substantive political/organizational maintenance model. Because these interaction terms were not found to be significant, I omitted them from the findings that I present.

TABLE 3  
UNSTANDARDIZED OLS REGRESSION ESTIMATES FOR  
SENTENCE LENGTH MODEL OMITTING THE  
ORGANIZATIONAL CONTEXT

Variable	Sentence Length Model
Legal:	
Severity .. ...	8.56* (.55)
History ... ..	5.51* (.52)
Weapon .....	9.93* (2.75)
Multiple charges .....	3.94* (1.75)
Processing:	
Plea .. ...	-12.17* (2.67)
Attorney type.....	1.36 (3.01)
Social status:	
Race .....	1.91 (1.92)
N .....	300
R <sup>2</sup> .....	.61

NOTE.—Values in parentheses are SEs.

\* $P < .05$

nificantly increase the probability of receiving a prison term. As the organizational maintenance theory would predict, cases that plead guilty also have a lower probability of receiving a prison term than cases with a trial disposition. Similarly, when I regress sentence length on racial, legal, and processing variables without reference to the organizational context, the legal variables and plea are statistically significant predictors of sentence length. Increases in offense severity, history, charges, and the use of a weapon significantly increase the number of months sentenced to prison, while a plea of guilty significantly decreases the number of months.

Next, I ask whether the organizational maintenance theory is supported across organizational contexts. If legal and processing variables predict prison and sentencing regardless of organizational context, then the organizational maintenance theory should be supported regardless of the organizational environment in which a case is processed. If, however, the variables that predict prison and sentencing length vary across organizational contexts, the type of sentencing theory supported should depend on the organizational context. To test whether the effects of vari-

ables on sentencing vary across organizational contexts, I split the sample according to the level of bureaucratization of judicial and prosecutorial activities using two dimensions of bureaucratization: the complexity of the division of labor and the decentralization of decision making. I estimate the effects of racial, legal, and processing variables on prison under conditions of low and high judicial and prosecutorial bureaucratization (see table 4) and on sentence length under conditions of low and high judicial and prosecutorial bureaucratization (see table 5). The findings from the analysis of both judicial and prosecutorial contexts reveal that the formal legal theory is supported under conditions of low bureaucratization and the organizational maintenance theory is supported under conditions of high bureaucratization, whether it be judicial or prosecutorial. This pattern persists for both the prison and the sentence length decisions, but is stronger for the sentence length decision.

#### The Organizational Context of the Prison Decision

As hypothesized, the organizational context conditions which sentencing theory explains whether a defendant is given a prison sentence. When the sample is split, there is support for the formal legal theory of sentencing under conditions of low bureaucracy and support for the organizational maintenance theory under conditions of high bureaucracy. This difference persists for both the judicial and the prosecutorial context of the imprisonment decision.

*Judicial bureaucratization and the prison decision.*—When cases are processed in a court with *low judicial bureaucratization*—low complexity in the judicial division of labor and low decentralization in judicial decision making—the organization of the prison decision corroborates the formal legal theory of sentencing. The estimates for the variables in equations (1a) and (1c) in table 4 show that the coefficients for the legal variables are statistically significant while the coefficients for the processing and race variables are insignificant when cases are processed in courts where the work of judges is organized with noncomplex mixed dockets or centralized individual calendars. The coefficients for three of the legal variables, severity, history, and weapon, are statistically significant in the predicted direction: the greater the offense severity, the greater the probability of prison; the greater the prior criminal history, the greater the probability of prison; cases in which a weapon is used have a higher probability of prison. Since the dispositional aspect of the sentencing decision in this setting reflects the formal legal model of sentencing, race and plea have no significant effect on the imprisonment decision.

In contrast, when cases are processed in a court with *high judicial bureaucratization*—high complexity in the judicial division of labor and



TABLE 4

## UNSTANDARDIZED LOGIT ESTIMATES FOR PRISON MODELS FOR VARIOUS ORGANIZATIONAL CONTEXTS

VARIABLE	JUDICIAL BUREAUCRATIZATION				PROSECUTORIAL BUREAUCRATIZATION			
	Complexity of Division of Labor		Decentralization of Decision Making		Complexity of Division of Labor		Decentralization of Decision Making	
	Low (1a)	High (1b)	Low (1c)	High (1d)	Low (2a)	High (2b)	Low (2c)	High (2d)
Legal:								
Severity.....	.76* (.11)	.68* (.09)	.79* (.13)	.68* (.08)	.84* (.14)	.66* (.08)	1.27* (.23)	.63** (.07)
History.....	1.37* (.13)	1.10** (.09)	1.31* (.14)	1.18* (.09)	1.43* (.16)	1.14** (.08)	1.77* (.27)	1.16** (.08)
Weapon .....	1.61* (.64)	2.10* (.48)	1.84* (.72)	1.96* (.44)	.56 (.85)	2.43** (.46)	.83 (1.02)	2.16** (.42)
Multiple charges. .	.18 (.32)	-.05 (.27)	-.02 (.36)	.12 (.25)	.02 (.37)	.06 (.25)	-1.02 (.58)	.23* (.22)
Processing:								
Plea. ....	-.61 (.90)	-1.64* (.52)	-.70 (1.01)	-1.48* (.48)	-.48 (1.06)	-1.63* (.50)	-1.06 (2.18)	-1.38* (.45)
Attorney type ....	-.54 (.58)	-.66 (.36)	-1.17 (.81)	-.44 (.34)	-.02 (.61)	-.79* (.36)	-1.89 (1.30)	-.45 (.32)
Social status:								
Race.....	-.55 (.46)	-.08 (.28)	-.57 (.58)	-.31 (.26)	-.39 (.52)	-.15 (.26)	.71 (1.14)	-.37 (.24)
N .....	701	745	403	953	503	943	314	1,132
-2 log L .....	348*	454*	256*	553*	249*	564*	180*	639*
df.....	7	7	7	7	7	7	7	7

NOTE.—Values in parentheses are SEs.

\* Statistically significant difference exists between low and high categories.

\*  $P < .05$

TABLE 5

UNSTANDARDIZED OLS REGRESSION ESTIMATES FOR SENTENCE LENGTH MODELS FOR VARIOUS ORGANIZATIONAL CONTEXTS

VARIABLE	JUDICIAL BUREAUCRATIZATION				PROSECUTORIAL BUREAUCRATIZATION			
	Complexity of Division of Labor		Decentralization of Decision Making		Complexity of Division of Labor		Decentralization of Decision Making	
	Low (1a)	High (1b)	Low (1c)	High (1d)	Low (2a)	High (2b)	Low (2c)	High (2d)
<b>Legal:</b>								
Severity .....	8.06* (.93)	8.83* (.71)	8.41* (1.11)	8.56* (.65)	8.22* (1.22)	8.63* (.64)	9.82* (1.42)	8.36* (.62)
History . . . . .	6.23* (.86)	5.21* (.66)	7.03* (.98)	4.92* (.62)	6.81* (1.14)	5.14** (.60)	6.32* (1.45)	5.41* (.56)
Weapon.....	7.36 (5.07)	10.19* (3.36)	6.79 (5.54)	10.44* (3.18)	10.02 (7.11)	9.36* (3.03)	-4.97 (6.56)	12.21** (3.05)
Multiple charges...	2.80 (2.76)	3.22 (2.32)	4.38 (3.26)	2.64 (2.09)	3.80 (3.63)	3.27 (2.04)	5.68 (3.94)	3.60 (1.94)
<b>Processing:</b>								
Plea.. . . . .	1.22 (6.24)	-14.29** (3.11)	1.32 (6.74)	-14.68** (2.96)	(.82) (8.53)	-13.65** (2.89)	-22.63* (8.61)	-11.32** (2.86)
Attorney type .....	4.36 (6.09)	-78 (3.57)	4.35 (9.12)	.03 (3.24)	3.90 (6.91)	.66 (3.37)	17.13 (9.59)	.10 (3.22)
<b>Social status:</b>								
Race.....	-1.01 (4.28)	3.38 (2.37)	.39 (5.21)	2.42 (2.18)	-.48 (5.81)	1.97 (2.12)	-3.18 (11.07)	2.42 (2.06)
N.....	113	187	89	211	81	219	49	251
R <sup>2</sup> .....	.53	.65	.55	.64	.55	.64	.67	.62

NOTE.—Values in parentheses are SEs.

\* Statistically significant difference exists between low and high categories.

\*  $P < .05$

high judicial decentralization in decision making—plea dispositions as well as legal factors affect the prison decision and the organizational maintenance theory of sentencing is substantiated. The estimates for equations (1b) and (1d) in table 4 demonstrate that three legal variables and the plea variable are statistically significant when the work of the judges is organized with complex specialized dockets and highly decentralized master calendaring systems. In this setting the probability of prison increases as criminal history and severity increases; the probability of prison is greater for cases with a weapon; and the probability of prison is greater for cases that plead guilty. Inasmuch as the prison decision in this setting reflects organizational maintenance rationality, race has no significant effect on the imprisonment decision.

*Prosecutorial bureaucratization and the prison decision.*—The models predicting prison under conditions of low and high prosecutorial bureaucratization show results similar to those under conditions of low and high judicial bureaucratization. When cases are processed in a court with *low prosecutorial bureaucratization*—low complexity in the prosecutorial division of labor and low decentralization in prosecutorial decision making—the organization of the imprisonment decision verifies the formal legal theory of sentencing. As the coefficients for the variables in equations (2a) and (2c) in table 4 show, legal variables are the statistically significant predictors of imprisonment for cases in courts where there is a noncomplex prosecutorial division of labor because of the small number of divisions or where prosecutorial decision making is centralized in the hands of the chief prosecutor; as offense severity or criminal history increases, the probability of receiving a prison sentence significantly increases. Because imprisonment decisions in this setting simulate formal legal rationality, neither plea nor race have a significant effect on the probability of imprisonment.

Conversely, when cases are processed in a court with *high prosecutorial bureaucratization*—high complexity in the prosecutorial division of labor and high decentralization in prosecutorial decision making—the organization of the prison decision supports the organizational maintenance theory of sentencing. The coefficients for the variables in equations (2b) and (2d) in table 4 reveal that both legal and processing variables are significant predictors of the prison decision when cases are processed in courts where the prosecutor's office is characterized by a large number of divisions or where prosecutorial decision making is decentralized such that assistant prosecutors engage in plea bargaining. When the complexity of the prosecutorial division of labor is high (see eq. [2b] in table 4), the probability of prison is significantly related to offense severity, criminal history, weapon, plea, and attorney type; as offense severity or criminal history increases, the probability of prison increases; the probability of

prison decreases for cases involving a weapon, cases entering a guilty plea, and cases represented by a private attorney. Similarly, when the decentralization of prosecutorial decision making is high (see eq. [2d] in table 4), the probability of prison significantly increases with offense severity, criminal history, and the possession of a weapon, while it significantly decreases for cases with a guilty plea. Race has no influence on imprisonment decisions in this setting.

Legal variables are significant in predicting the imprisonment decision regardless of the organizational context. While the effect of plea is not significant in predicting prison for the low-bureaucratization models and is significant in predicting prison for the high-bureaucratization models for judicial (eqq. [1b] and [1d] in table 4) and prosecutorial contexts (eqq. [2b] and [2d] in table 4), it is noteworthy to add that there is no statistically significant *difference* between the effect of plea when comparing the magnitude of the effect of plea in equations (1b) and (1d) and when comparing the magnitude of the effect of plea in equations (2b) and (2d). A *t*-test is used to compare the magnitude of the difference between each variable in the low- and high-bureaucratic contexts since it takes into account the differences in the SEs and the number of cases in each subsample. Although the magnitude of the difference between the plea effect in the low-bureaucratic context and in the high-bureaucratic context is not statistically significant, the magnitude of the plea coefficient predicting prison is always greater under conditions of high bureaucratization. When we analyze the difference between the magnitudes of the plea effect under conditions of low and of high bureaucratization for the sentence length decision in the next section, the plea effect is *also* always greater under conditions of high bureaucratization. Furthermore, the difference between the magnitudes of the plea effect under conditions of low and of high bureaucratization for the sentence length decision is statistically significant.<sup>20</sup> This suggests that the organizational context affects the sentence length decision even more than it affects the imprisonment decision.

### The Organizational Context of the Sentence Length Decision

The results from the models predicting the sentence length decision show many similarities to the results of the models predicting the imprisonment decision. The organizational context determines which sentencing theory best predicts the number of months of prison given to those individuals

<sup>20</sup> While not the emphasis of this article, it is interesting to observe that the effects of offense severity and criminal history significantly increase, while the effects of weapon significantly decrease, as judicial bureaucratization increases.

with a prison sentence. When estimations of the determinants of the sentence length decision are made for cases processed in lowly and in highly bureaucratized contexts, the formal legal theory is favored for cases processed in lowly bureaucratized contexts and the organizational maintenance theory is corroborated for cases processed in highly bureaucratized contexts. This pattern is highly stable for the judicial and prosecutorial contexts.

*Judicial bureaucratization and the sentence length decision.*—When an offender is sentenced in a court with low judicial bureaucratization, the organization of the sentence length decision mirrors the formal legal theory of sentencing. As the estimates for the variables predicting sentence length in equations (1a) and (1c) in table 5 illustrate, legal variables are statistically significant and processing and race variables are statistically insignificant in predicting the sentence length decision when there are noncomplex mixed dockets or centralized individual calendars. The two legal variables that predict the number of months of imprisonment are offense severity and criminal history; as the offense severity or the criminal history of an offender increases, the number of months of prison in the sentence increases. Neither plea nor race are significant predictors of sentence length in this setting, where formal legal rationality predominates.

Conversely, the sentence length decision for offenders processed in courts with high judicial bureaucratization mirrors organizational maintenance processes. As the results from equations (1b) and (1d) in table 5 testify, the sentence length decision for cases processed in courts with high complexity (specialized dockets) in the judicial division of labor or high decentralization in judicial decision making (master calendars) are significantly affected by legal and processing variables. The effects of the legal variables on sentence length are positive, and the effect of the plea variable is negative and very strong; as offense severity or criminal history increases, the number of months of prison increases; cases that plead guilty receive fewer months of prison than cases that go to trial. The formal legal rationality of the sentence length decision featured in this setting obviates any significant effect of either plea or race.

*Prosecutorial bureaucratization and the sentence length decision.*—The results from the models predicting sentence length under conditions of low and high prosecutorial bureaucratization closely resemble the results from the models predicting sentence length under conditions of low and of high judicial bureaucratization.

The sentence length decision for cases processed in courts with low prosecutorial bureaucratization substantiates the formal legal theory of sentencing. As the results from equation (2a) in table 5 denote, the significant predictors of sentence length are legal variables when cases are

processed in courts where the prosecutor's office has a small number of divisions or where the plea bargaining decisions are centralized in the hands of the chief prosecutor. The only aberrant pattern in the analysis is found in the results from equation (2c) in table 5, which shows that the organizational maintenance theory is supported when the decentralization of prosecutorial decision making is low. In this case, both legal variables and plea are statistically significant predictors of sentence length, with the effect of the plea variable having a greater negative effect in the lowly bureaucratized context: as offense severity or criminal history increases, the number of months in prison increases; pleading guilty decreases the number of months in prison.

For cases processed in a court with high prosecutorial bureaucratization, the sentence length decision corroborates the organizational maintenance theory of sentencing. Findings from the model predicting an offender's sentence length when the prosecutors office is large and complex ([2b] in table 5), as well as findings from the model in which the plea bargaining is decentralized in the hands of assistant prosecutors ([2d] in table 5), indicate that legal variables as well as the processing variable of plea influence the sentence length decision: the number of months in the prison sentence increases with offense severity or criminal history; the number of months in the prison sentence is less for cases that plead guilty rather than going to trial. Because the sentence length decision in this setting reflects an organizational maintenance rationality, there is no effect of race on the sentence length decision.

The absence of a significant plea effect on sentence length in most of the low-bureaucratization conditions, coupled with the presence of a significant plea effect in all of the high-bureaucratization conditions, suggests that bureaucratic administrative structures designed to make case processing more efficient has an effect on the organization of sentencing. A defendant who responds to incentive systems to plead guilty in bureaucratically organized judicial and prosecutorial contexts is less likely to be sent to prison and more likely to receive a lenient prison term than a defendant pleading guilty in a less bureaucratically organized court. The statistically significant difference in the effects of plea on sentence length across all comparisons of low and high bureaucratization is particularly persuasive for the argument that organizational contexts condition the organization of sentencing for both the prison and sentence length aspects of the sentencing decision.

## CONCLUSIONS

These results demonstrate the strengths and limitations of each of the major perspectives on criminal sentencing by indicating the organiza-

tional context in which each theoretical perspectives applies. It is not surprising that the effects of legal variables are significant and strong regardless of the organizational context; this finding suggests that the formal legal perspective on sentencing must be considered in all court contexts. The processing measure of plea significantly affects sentencing in particular organizational contexts in a manner consistent with the precepts of the organizational maintenance theory. When compared with the results of previous research, these findings suggest the need to develop theories and empirical models that consider and directly measure the organizational as well as other contexts in which sentencing takes place.

The formal legal theory receives the type of support under conditions of low judicial and prosecutorial bureaucratization that it often receives in the noncontextual research. Verification of the formal legal theory and negation of the organizational maintenance, substantive political, and combined substantive political/organizational maintenance theories is consistently shown in the low-bureaucratization models of sentencing, where legal variables are the primary determinants of imprisonment and sentence length. The fact that neither plea nor race affects sentencing in counties with low levels of judicial and prosecutorial bureaucratization indicates that an institutionalized system of sentence reduction based on race discrimination or plea bargaining is absent in this organizational context. The findings for the models predicting imprisonment and sentence length under conditions of low bureaucratization correspond to early noncontextual studies finding positive effects of legal variables and no effect of either processing or racial variables on sentencing disposition and duration. However, the limitations of the formal legal theory are shown in the models predicting sentence type and length under conditions of high judicial and prosecutorial bureaucratization. Using a sample from a sentencing guidelines state favors the formal legal theory in all contexts; state policies endorsing uniformity and the exclusion of social and processing considerations in sentencing determinations increases the importance of legal factors and minimizes the probability that race and plea will effect sentencing in any context. Hence, effects of race and plea in any organizational context in a sentencing guidelines state would provide strong evidence for the substantive political perspective and the organizational maintenance perspectives, respectively.

The validity of the organizational maintenance theory is evident in the models predicting imprisonment and length under conditions of high bureaucratization in judicial and prosecutorial administrative structures. The strong negative effect of plea, coupled with the positive effect of legal variables, in these models is consistent with the empirical sentencing literature, arguing that sentencing processes reflect organizational maintenance processes. The evidence from this research shows that sentence

reduction for pleading guilty is institutionalized in courts where the work of prosecutors and judges is bureaucratically structured. What is most notable is the consistent patterns for the plea effect across the various dimensions of prosecutorial bureaucracy; plea is consistently insignificant in the low-bureaucratization models and consistently significant in the high-bureaucratization models. While the organizational maintenance theory is substantiated under conditions of high bureaucratization, its lack of predictive ability under conditions of low judicial and prosecutorial bureaucratization reveal its limitations.

Although findings supporting the substantive political or the combined substantive political/organizational maintenance model of sentencing are lacking in this research, this may be the result of several factors associated with Minnesota. First, the racial composition of the general and the criminal population in Minnesota is largely white. This racial composition probably affects the relationship between race and sentencing; further research should extend the newly emerging research agenda on contextual effects by examining the results from my model in states that have different racial compositions. For example, when my findings are compared to those of Myers and Talarico (1987), I do not find the racial effects that they report. This is not surprising, however, since their data comes from a southern nonguidelines state where there is a predominately minority criminal population with a greater racial mix in the general population. A second characteristic that distinguishes Minnesota is the sentencing guidelines policy; this policy favors the formal legal theory of sentencing and, hence, affects the relationship between race and sentencing. The lack of a significant relationship between race and sentencing in my analysis may not be found in studies of sentencing in states without sentencing guidelines. Moreover, the absence of racial effects for sentencing in Minnesota does not imply that there are no effects of race on other dispositions, since many have argued that there is a hydraulic displacement of racial discretion from sentencing to plea bargaining in sentencing guidelines states. Differences in the effects of race on plea bargaining and sentencing across states with differing sentencing policies only reinforces the premise of this research that sentencing processes vary across settings.

A serious consideration of the organizational context perspective on sentencing reveals that what appear to be competing theories of sentencing are actually complementary theories that are capable of predicting the organization of sentencing in particular contexts. The consistent finding that legal variables related to offense severity and criminal history are important for predicting sentencing under conditions both of low and of high bureaucracy, while plea dispositions take on an added importance in the high-bureaucracy context, reveals the importance of the judicial



and prosecutorial context in structuring sentencing. Comparisons of sentencing in courts with low and with high levels of bureaucracy fail to support Tepperman's supposition that bureaucratically organized courts render sentencing processes that are less affected by legally irrelevant factors than nonbureaucratically organized courts. In fact, the results are more consistent with Eisenstein et al. (1988), who argue that the greater division of labor and decentralized decision making in bureaucratic courts creates loosely coupled unstable work groups that use individual discretion such that legally irrelevant factors influence sentencing. While my findings are consistent with the theories of Eisenstein et al. (1988), they are in no way a direct test since I offer no direct measure of either the informal relations or tightness of coupling among court actors in the bureaucratic and nonbureaucratic contexts. It is quite possible that tight coupling among court actors only occurs if the court actors are in contexts in which the organization of work is nonbureaucratic and they are regular participants in a local legal and community culture. The assumption that courts in rural areas are nonbureaucratized and involve tight coupling among court actors engaging in other informal relationships in the community merits further investigation since rural courts often share personnel in a manner that does not ensure that the judge hearing a case in a particular court belongs to a stable local legal culture or even the local community culture.

What are the consequences of the variation in sentencing processes across courts with different organizational structures in a sentencing guidelines state as evidenced in this study? Criminal justice professionals have attempted to establish an objective categorical system of sentencing to guard against the introduction of legally irrelevant criteria in sentencing decisions concerning imprisonment and sentence length. My results should not be seen as an indictment of attempts to improve sentencing consistency. On the contrary, the explicit criteria set forth in the Minnesota sentencing guidelines may have attenuated the effects of race on sentencing outcomes. The intent of this study is to demonstrate that, even with carefully drawn standards, sentencing will vary across courts that vary in their formal and informal organization. Professions seek legitimacy by demonstrating the objective nature of their decision making. This trend toward seeking consistency and fairness in sentencing is admirable, but a false sense of confidence in objective measures is created if we ignore the local context of sentencing. I hope that I have shown that it is premature to close the issue of the effects of legal, social, and processing factors on sentencing decisions until the contextual effects of the organizational, as well as social, environments of courts have been examined.

The findings of this research point to the limitations of any sentencing

theory that fails to address the organizational context of sentencing. Future theoretical developments and research designs need to take seriously the organizational context perspective of sentencing. Similarly, my conclusions suggest the need for including the racial mix and state sentencing policies of courts where cases are processed. The organizational context is only one of several important contexts that influence the role of legal, social status, and processing factors in sentencing decisions.

At the theoretical level, it is important to synthesize the previous theories of sentencing such that each is seen as one of several ways that sentencing processes are organized. As sentencing research continues to examine sentencing in more contexts, the need to determine the context in which each of the current theories serves as the best model becomes more imperative. Since my results show that some theories are better suited for explaining sentencing when judicial and prosecutorial structures are lowly bureaucratized while others are best at explaining sentencing when structures are highly bureaucratized, researchers should not omit the organizational context of sentencing. The complexity and variability of sentencing processes, however, will not be fully captured until research expands on the model in this paper to examine the organization of sentencing in jurisdictions that vary in their organizational, political, policy, and social contexts.

TABLE A1  
CORRELATION MATRIX FOR PRISON MODEL

[illegible]

## REFERENCES

- Austin, Thomas. 1981. "The Influence of Court Location on Type of Criminal Sentence. The Rural-Urban Factor." *Journal of Criminal Justice* 9:305-16.
- Balbus, Isaac D. 1977. *The Dialectics of Legal Repression*. New Brunswick, N.J.: Transaction.
- Benson, Michael, and Esteban Walker. 1988. "Sentencing the White-Collar Offender." *American Sociological Review* 53:294-302.
- Bernstein, Ilene Nagel, William Kelly, and Patricia Doyle. 1977. "Societal Reaction to Deviants: The Case of Criminal Defendants." *American Sociological Review* 42:743-55.
- Blau, Peter. 1974. *On the Nature of Organizations*. New York: Wiley.
- Blumberg, Abraham. 1967. "The Practice of Law as a Confidence Game." *Law and Society Review* 1:15-39.
- Brereton, David, and Jonathan D. Casper. 1982. "Does It Pay to Plead Guilty? Differential Sentencing and the Function of Criminal Courts." *Law and Society Review* 16:45-70.
- Burke, Peter, and Austin Turk. 1975. "Factors Affecting Postarrest Dispositions: A Model for Analysis." *Social Problems* 22:313-32.
- Chambliss, William, and Robert Seidman. 1971. *Law, Order and Power*. Reading, Mass.: Addison-Wesley.
- Champion, Dean, ed. 1989. *The U.S. Sentencing Guidelines*. New York: Praeger.
- Chiricos, Theodore G., and Gordon P. Waldo. 1975. "Socioeconomic Status and Criminal Sentencing: An Empirical Assessment of a Conflict Perspective." *American Sociological Review* 40:753-72.
- Clarke, Stevens H., and Gary G. Koch. 1977. "Alaska Felony Sentencing Patterns. A Multivariate Statistical Analysis." Anchorage: Alaska Judicial Council.
- Eisenstein, James, Roy Flemming, and Peter Nardulli. 1988. *The Contours of Justice: Communities and Their Courts*. Boston: Little, Brown.
- Eisenstein, James, and Herbert Jacob. 1977. *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Little, Brown.
- Farrell, Ronald, and Malcolm Holmes. 1991. "The Social and Cognitive Structure of Legal Decision-Making." *Sociological Quarterly* 32:529-42.
- Flemming, Roy, Peter Nardulli, and James Eisenstein. 1992. *The Craft of Justice: Politics and Work in Criminal Court Communities*. Philadelphia: University of Pennsylvania Press.
- Garfinkel, Harold. 1949. "Research Note on Inter- and Intra-Racial Homicides." *Social Forces* 27:369-81.
- Greenberg, David, and Drew Humphries. 1980. "The Co-optation of Sentencing Guidelines." *Crime and Delinquency* 26:206-25.
- Greenwood, Peter, Joan Petersilia, and Franklin Zimring. 1980. *Age, Crime and Sanctions: The Transition from Juvenile to Adult Court*. Report R-2642-NIJ. Santa Monica, Calif.: RAND.
- Haas, Eugene, Richard Hall, and Norman Johnson. 1966. "Toward an Empirically Derived Taxonomy of Organizations." In *Studies on Behavior in Organizations*, edited by Raymond V. Bowers. Athens: University of Georgia Press.
- Hagan, John. 1974. "Extra-Legal Attributes and Criminal Sentencing: An Assessment of a Sociological Viewpoint." *Law and Society Review* 8:357-83.
- . 1975. "Parameters of Criminal Prosecution: An Application of Path Analysis to a Problem of Criminal Justice." *Journal of Criminal Law, Criminology and Police Science* 65:536-44.
- . 1977. "Criminal Justice in Rural and Urban Communities: A Study of the Bureaucratization of Justice." *Social Forces* 55:597-612.
- Hagan, John, and Kristin Bumiller. 1983. "Making Sense of Sentencing: A Review and Critique of Sentencing Research." Pp. 1-54 in *Research in Sentencing: The*

- Search for Reform*, edited by A. Blumstein, J. Cohen, S. E. Martin, and M. H. Tonry. Washington, D.C.: National Academy of Science.
- Hagan, John, John Hewitt, and Duane Alwin. 1979. "Ceremonial Justice: Crime and Punishment in a Loosely Coupled System." *Social Forces* 58:506-27.
- Hagan, John, Ilene Nagel, and Celesta Albonetti. 1980. "The Differential Sentencing of White-Collar Offenders in Ten Federal District Courts." *American Sociological Review* 45:802-20.
- Hage, Gerald, and Michael Aiken. 1969. "Routine Technology, Social Structure and Organizational Goals." *Administrative Science Quarterly* 14:366-77.
- Hall, Richard, J. Eugene Haas, and Norman Johnson. 1967. "Organizational Size, Complexity and Formalization." *American Sociological Review* 32:903-12.
- Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153-61.
- Heumann, Milton. 1975. "A Note on Plea Bargaining and Caseload Pressure." *Law and Society Review* 9:515-28.
- . 1978. *Plea Bargaining: The Experiences of Prosecutors, Judges and Defense Attorneys*. Chicago: University of Chicago Press.
- Heumann, Milton, and Colin Loftin. 1979. "Mandatory Sentencing and the Abolition of Plea Bargaining: The Michigan Felon Firearms Statute." *Law and Society Review* 13:393-430.
- Holmes, Malcolm, Howard Daudistel, and William Taggart. 1992. "Plea Bargaining Policy and State District Court Caseloads. An Interrupted Time Series Analysis." *Law and Society Review* 26:139-59.
- Johnson, Elmer. 1957. "Selective Forces in Capital Punishment." *Social Forces* 36:165-69.
- Kleck, Gary. 1981. "Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence with Additional Evidence on the Death Penalty." *American Sociological Review* 46:783-805.
- Knapp, Kay. 1982. "Impact of the Minnesota Sentencing Guidelines on Sentencing Structures." *Hamline Law Review* 5:237-70.
- LaFree, Gary. 1980. "The Effect of Sexual Stratification by Race on Official Reactions to Rape." *American Sociological Review* 45:842-54.
- . 1985. "Adversarial and Nonadversarial Justice: A Comparison of Guilty Pleas and Trials." *Criminology* 23:289-312.
- . 1989. *Rape and Criminal Justice: The Social Construction of Sexual Assault*. Belmont, Calif.: Wadsworth.
- Levin, Martin. 1977. *Urban Politics and the Criminal Courts*. Chicago: University of Chicago Press.
- Lizotte, Alan J. 1978. "Extra-Legal Factors in Chicago's Criminal Courts: Testing the Conflict Model of Criminal Justice." *Social Problems* 25:564-80.
- Lukács, Georg. 1971. *History and Class Consciousness*. Cambridge, Mass.: MIT Press.
- Luskin, Mary Lee, and Robert Luskin. 1986. "Why So Fast? Why So Slow? Explaining Case Processing Time." *Journal of Criminal Law and Criminology* 77:190-214.
- Martin, Roscoe. 1934. *The Defendant and Criminal Justice*. Bulletin no. 34-37. University of Texas, Bureau of Research in the Social Sciences.
- Martin, Susan. 1983. "The Politics of Sentencing Reform: Sentencing Guidelines in Pennsylvania and Minnesota." Pp. 265-304 in *Research in Sentencing: The Search For Reform*, edited by A. Blumstein, J. Cohen, S. Marti, and M. H. Tonry. Washington, D.C.: National Academy of Science.
- Marx, Karl. 1859. *A Contribution to the Critique of Political Economy*. New York: Vintage.
- Meeker, James, and Henry Pontell. 1985. "Court Caseloads, Plea Bargains and Crim-

- inal Sanctions: The Effects of Section 17 P.C. in California." *Criminology* 23: 119-43.
- Michels, Robert. 1915. *Political Parties*. Glencoe, Ill.: Free Press.
- Miethe, Terrance. 1987. "Charging and Plea Bargaining Practices under Determinant Sentencing: An Investigation of the Hydraulic Displacement of Discretion." *Journal of Criminal Law and Criminology* 78:101-22.
- Miethe, Terrance, and Charles Moore. 1985. "Socioeconomic Disparities under Determinant Sentencing Systems. A Comparison of the Preguideline and Postguideline Practices in Minnesota." *Criminology* 23:337-63
- Minnesota Sentencing Guidelines Commission. 1984 *The Impact of the Minnesota Sentencing Guidelines Commission: Three Year Evaluation*. Saint Paul: Minnesota Sentencing Guidelines Commission.
- Moore, Charles, and Terrance Miethe. 1986. "Regulated and Unregulated Sentencing Decisions: An Analysis of First-Year Practices under Minnesota's Felony Sentencing Guidelines." *Law and Society Review* 20:253-77.
- Myers, Martha, and Gary LaFree. 1982. "Sexual Assault and Its Prosecution: A Comparison with Other Crimes " *Journal of Criminal Law and Criminology* 73: 1282-1305.
- Myers, Martha, and Susan Talarico. 1986. "Urban Justice, Rural Injustice: Urbanization and Its Effect on Sentencing." *Criminology* 24:367-91
- . 1987. *The Social Contexts of Criminal Sentencing*. New York: Springer.
- Nardulli, Peter. 1979. *The Courtroom Elite*. New York: Ballinger.
- Nardulli, Peter, James Eisenstein, and Roy Flemming. 1988. *The Tenor of Justice*. Urbana: University of Illinois Press.
- Newman, Donald J. 1956. "Pleading Guilty for Considerations." *Journal of Criminal Law, Criminology and Police Science* 46:780-90
- Nonet, Philippe, and Philip Selznick. 1978. *Law and Society in Transition: Toward Responsive Law*. New York: Octagon.
- Pope, Carl. 1976. "The Influence of Social and Legal Factors on Sentencing Dispositions: A Preliminary Analysis of Offender Based Transaction Statistics." *Journal of Criminal Justice* 4:203-21.
- Pugh, Derek, David Hickson, C. R. Hinings, and C. Turner. 1968. "Dimensions of Organizational Structure." *Administrative Science Quarterly* 13:65-105.
- Quinney, Richard. 1970. *The Social Reality of Crime*. Boston: Little, Brown.
- Savelsberg, Joachim. 1992. "Law That Does Not Fit Society: Sentencing Guidelines and Substantivized Law." *American Journal of Sociology* 97:1346-81.
- Sellin, Thorsten. 1938. *Culture, Conflict and Crime*. New York: Social Science Research Council.
- Selznick, Philip. 1966. *TVA and the Grass Roots*. New York: Harper & Row.
- Spohn, Cassia, John Gruhl, and Susan Welch. 1982. "The Effect of Race on Sentencing: A Re-Examination of an Unsettled Question " *Law and Society Review* 16: 71-88
- Sudnow, David. 1965. "Normal Crimes: Sociological Features of the Penal Code in a Public Defender Office." *Social Problems* 12:255-76.
- Tepperman, L. 1973. "The Effects of Court Size on Organization and Procedure." *Canadian Review of Sociology and Anthropology* 10 346-65.
- Thomson, Randall J., and Mathew T. Zingraff. 1981. "Detecting Sentence Disparity: Some Problems and Evidence." *American Journal of Sociology* 86:869-80.
- Uhlman, Thomas, and Darlene Walker. 1980. "He Takes Some of My Time; I Take Some of His: An Analysis of Judicial Sentencing Patterns in Jury Cases." *Law and Society Review* 14:323-41.
- Unnever, James D., Charles Frazier, and John C. Henretta 1980. "Race Differences in Criminal Sentencing." *Sociological Quarterly* 21:197-207.

- von Hirsch, Andrew, Kay Knapp, and Michael Tonry. 1987. *The Sentencing Commission and Its Guidelines*. Boston: Northeastern University Press.
- Warriner, Charles, Richard Hall, and Bill McKelvey. 1981. "The Comparative Description of Organizations: A Research Note and Invitation." *Organizational Studies* 2:100-120.
- Weber, Max. 1954. *Law in Economy and Society*. Edited by Max Rheinstein and translated by Edward Shils and Max Rheinstein. Cambridge, Mass.: Harvard University Press.
- Wheeler, Stanton, David Wiseburd, and Nancy Bode. 1982. "Sentencing the White-Collar Offender: Rhetoric and Reality." *American Sociological Review* 47:641-59.
- Wilbanks, William. 1987. *The Myth of a Racist Criminal Justice System*. Monterey, Calif.: Brooks/Cole.
- Wooldredge, John. 1989. "An Aggregate-Level Examination of the Caseload Pressure Hypothesis." *Journal of Quantitative Criminology* 5:259-83.
- Zatz, Marjorie S. 1985. "Pleas, Priors and Prison: Racial/Ethnic Differences in Sentencing." *Social Science Research* 14:169-93.
- Zatz, Marjorie, and John Hagan. 1985. "Crime, Time and Punishment: An Exploration of Selection Bias in Sentencing Research." *Journal of Quantitative Criminology* 1:103-26.

# Structural Sources of Socioeconomic Segregation in Brazilian Metropolitan Areas<sup>1</sup>

Edward E. Telles

*University of California, Los Angeles*

This article evaluates the impact of industrialization and urbanization on residential segregation by income among Brazilian metropolitan areas. Using data from the 1980 census of Brazil, the author finds that more-industrialized areas have lower segregation because they have lower income inequality. However, urbanization, particularly population size, explains most of the variation in segregation among metropolitan areas. The conclusion is that the extent of urbanization, which is independent of industrialization, is key to understanding socioeconomic spatial inequalities in the large and rapidly growing cities of less-developed countries, but that conclusions regarding the effects of industrialization through industrial location or investment in real estate have been overgeneralized.

On October 18, 1992, much of Rio de Janeiro's middle class felt besieged by the city's poor. Busloads of youths from poor communities in the "North Zone" paraded across Copacabana and Ipanema beaches in the "South Zone," startling the predominantly white, middle-class beachgoers into running away. In a few cases, the youths engaged in fistfights and in the petty theft of sandals, watches, and sunglasses, but the worst crime of all seemed to be their massive presence on the beach. The reactions to this event (*arrastão*) by South Zone residents revealed their prejudices and insecurities about the "poor, dark-skinned" residents of the North Zone and of the highly visible, but socially distant, *favelas* (housing on illegally occupied land) of nearby hillsides (Veja 1992; *Folha de São Paulo* 1992). The incident and the reaction to it manifested

<sup>1</sup> Research for this paper was supported by a UCLA Academic Senate Grant. I acknowledge the helpful comments of Eduardo Arriaga, Ken Bailey, Rebecca Emigh, Gary Hytrek, Martim Smolka, Roger Waldinger, Michael White, and the *AJS* reviewers. Send correspondence to Edward E. Telles, Department of Sociology, University of California, 264 Haines Hall, Los Angeles, California 90024-1551



deep-seated class and racial prejudices among the middle class throughout Brazil as they came to realize that the social and spatial barriers that had long kept class and racial groups apart were actually quite fragile.

In addition to aggravating tension among social classes, as this event dramatizes, segregation affects access by the poor to schools, jobs, health services, and public utilities and impairs the capacity of cities to contribute to economic and social development (Yujnovsky 1975; Angiotti 1993). As evidence for the United States suggests (Mayer and Jencks 1989; Massey and Denton 1993; Brooks-Gunn et al. 1993), segregation is also likely to impair child development in Brazil and elsewhere because it concentrates poverty into particular neighborhoods. On the other hand, a positive aspect is that segregation may promote local control and facilitate political mobilization among the poor and the working class (Smith 1979; Caldeira 1992). In metropolitan areas like São Paulo, class segregation has strengthened working-class identities, allowing Worker's Party activists to build an organizational base that links neighborhood groups with labor unions (Bava 1994).

This study seeks to understand segregation in Brazilian urban areas in the context of industrial and urban demographic change. In Brazil and throughout Latin America, uneven industrialization and growing social inequality have characterized the regional pattern of development in recent decades (Geisse and Sabatini 1988; Portes 1989). As these countries become predominantly urban, industrialization's effect on urban socioeconomic and spatial inequalities becomes increasingly important for understanding social development. Brazil presents an excellent setting for examining the effect of industrialization on segregation because of its high levels of income inequality and because industrialization is quite uneven across metropolitan areas, ranging from highly industrialized ones like São Paulo to others that have grown without the benefit of industrialization and are consequently left with bloated urban informal sectors (Evans 1979; Merrick and Graham 1979). I also emphasize urbanization itself as important for understanding segregation in Brazil. The rapid growth of metropolitan areas from migration and the great variation in population size must be understood independent of industrialization and as critical factors in producing spatial inequalities.

## THEORY AND HYPOTHESES

In this section, I first examine four groups of theories that relate industrialization to segregation. I then examine the relation between segregation and demographic factors (migration and population size), and finally, I discuss the special case of the Brasília metropolitan area.

## Socioeconomic Segregation in Brazil

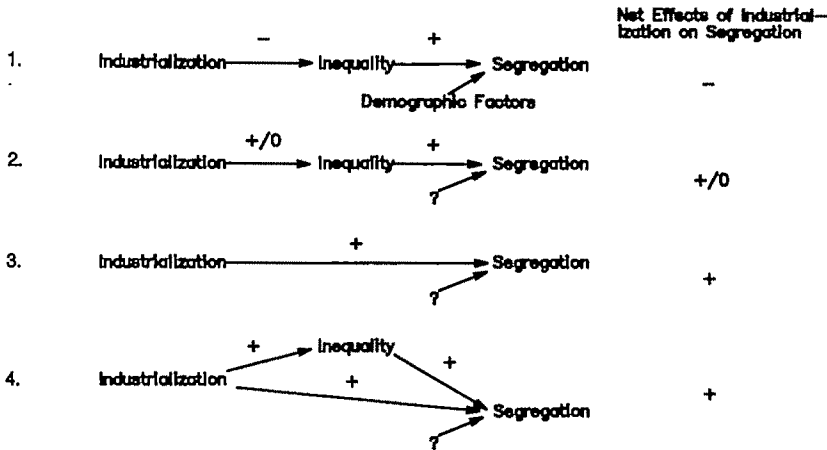


FIG. 1.—Hypothesized effects of industrialization on segregation

### Industrialization, Inequality, and Residential Segregation

Several theories about how industrialization affects segregation have been proposed, although the specifics often are not clearly stated. I try specifying the major relationships suggested by these theories and summarize them under four perspectives and hypotheses that I call (1) neo-classical, (2) Marxist, (3) industrial location and accumulation, and (4) underdeveloped industrialization. Clearly, some theoretical elements are found in more than one perspective, but the four are intended to characterize the major relationships between industrialization, inequality, and segregation. In all the perspectives that consider income inequality, I hypothesize that greater income inequality leads to greater segregation. These hypotheses are illustrated in figure 1 and discussed in the following paragraphs.

*Neoclassical perspective.*—Since the industrial revolution, the consistent and often large gains in per capita incomes accruing to industrializing societies have made industrialization the dominant criterion for development (Gilles et al. 1992). Presumably industrialization produces positive benefits for society. Wages increase, at least in the aggregate, and an increasingly large share of the labor force is incorporated into productive and higher-paid industrial jobs. Poor countries, like Brazil, have often sought to diminish their reliance on manufactured imports and have made industrial development a central goal in their attempts to overcome poverty and raise living standards (Portes and Benton 1984). As cities industrialize, neoclassical theorists would expect residential segregation

to fall because of decreasing income inequality and because lower income inequality would likely diminish differences in access to property and housing.

Nationally, neoclassical economic theory predicts that, with industrialization, national inequality will increase at first, as urban incomes increase while rural incomes remain stagnant. Later, as increasing numbers of the rural population migrate to cities in search of the better wages from industrialization, inequality will begin to decline, as fewer residents will be in poverty (Kuznets 1955). Within an urban area though, industrialization should thus consistently reduce inequality. Finally, I contend that segregation may be affected by population size and growth in which segregation is likely to be greater in larger areas and in those experiencing greater migration, which I will discuss following the discussion of alternative hypotheses about the effects of industrialization.

Thus, in light of neoclassical theory, I hypothesize:

**HYPOTHESIS 1.**—*Industrialization reduces segregation but only indirectly by reducing income inequality. In turn, segregation is reduced, as it is a spatial manifestation of income inequality. These effects occur along with demographic effects.*

**Marxist perspective.**—Inspired by Marx, other analysts claim that the industrial revolution polarized the population into classes and increased the division of labor, presumably leading to greater geographical segregation within cities (Engels [1845] 1973; Smith 1979). Based on evidence for recently industrialized countries, including Brazil, some development analysts challenge the neoclassical theory that industrialization eventually reduces income inequality (Chenery 1980). Rather, such effects may be country-specific and depend on specific development patterns. According to these analysts, Brazil's growth-oriented pattern has tended to produce greater inequality nationally compared to the equity-oriented, low-growth pattern illustrated by Sri Lanka or the rapid growth with equity pattern of Taiwan and Korea (Chenery 1980). Specifically, the benefits of industrialization for the working class have been insignificant compared to the disproportionate shares that have gone to industrial and financial capitalists (Kowarick and Ant 1988; Singer 1985). Strong military governments between 1964 and 1985 colluded with capitalists in maintaining high levels of inequality by keeping the minimum wage at very low levels, an amount sometimes paid even to workers in modern industrial jobs (Souza 1980; Singer 1985).<sup>2</sup> Under this theory, housing

<sup>2</sup> Much debate surrounded the magnitude of change in inequality in Brazil from 1960 to 1980, a period of rapid industrial growth. However, a thorough and careful evaluation by the World Bank of the empirical literature concluded that the structure of inequality in Brazil remained roughly the same or increased slightly from 1960 to

markets would similarly be more stratified, leading to greater segregation. Thus, this discussion suggests a counterhypothesis:

**HYPOTHESIS 2.**—*Residential segregation is unrelated to or increases with industrialization because preindustrial levels of inequality persist or inequality increases with industrialization.*

**Industrial location and accumulation perspective.**—Industrialization may have direct effects on the spatial distribution of urban populations. Residences tended to be adjacent to or part of the workplace in old industrial cities (Angiotti 1993). In large and complex metropolitan areas, the agglomeration of large-scale industries in key urban locations may drive the elites and middle class out (Schnore 1965) but might also attract higher-paid industrial or specialized workers to such locations, leaving the poor behind or perhaps displacing the poor of those areas (Logan and Molotch 1987; Fales and Moses 1972). Despite disagreement about which class groups are most affected by industrial location, both of these hypotheses indicate greater intraclass agglomeration and therefore greater socioeconomic segregation. On the other hand, such direct effects from industrialization might have been more important when transportation was less developed and workers lived close to their workplaces. The fact that manufacturing was based on large-scale production in even the most industrialized of Brazilian urban areas in 1980 permits us to test the theories of industrial location that may be irrelevant for the small-scale and flexible forms of industrialization that increasingly characterize modern economies (Sassen 1991).

In large cities that are highly class differentiated, the elites may attempt to segregate themselves and the middle class from the poor, as has been the case for São Paulo throughout most of this century (Caldeira 1992). It is not clear whether such *intent*s are universal or specific to particular cities. However, the *ability* of the urban elites and middle class to isolate themselves into income-homogeneous neighborhoods seems to depend on the capital that they can raise for building new housing tracts. Specifically, construction and real estate speculation may be intensified with greater capital accumulation drawn from industrial profits (Harvey 1985; Logan and Molotch 1987). This would lead to higher land prices, further stratifying residents spatially and raising the threshold of housing affordability for the poor. These theories tend to ignore the mediating effects that income distribution has on the relation between industrialization and segregation, and they would seem to support the following hypothesis:

1980, despite the growth (Pfefferman and Webb 1979). Wages in all sectors of the population increased during the period, although those at the very top increased the most. However, in the 1980s, a period of no growth, inequality increased though wages fell, to different degrees, at all levels (Bonelli and Sedlacek 1989).

**HYPOTHESIS 3.**—*Industrialization is directly related to greater residential segregation independent of its effects on inequality.*

*Underdeveloped industrialization perspective.*—Recently, some analysts have used the notion of “underdeveloped industrialization” to explain Brazil’s pattern of social development, inequality, and segregation (Ribeiro 1992; Kowarick and Campanário 1988). This perspective claims that, despite industrialization, working wages have declined, increasing income inequality. As the price of land rises, housing stratification increases. Specifically, capitalist industrialization in Brazil brings about greater control over salaries, more work hours, and worsening labor conditions and, thus, greater profits for industrialists. Lacking other alternatives for investment, industrialists invest a large portion of these profits in real estate, leading to higher land prices and thus greater income segregation because greater shares of the poor cannot afford housing of any kind (Geisse and Sabatini 1988). The failure of the state to control such activity and provide for the housing or general welfare of the poor allows such a process to go unchecked. Thus, industrialization, under this theory, would seem to negatively increase segregation both directly and indirectly via inequality.

**HYPOTHESIS 4.**—*Industrialization leads to greater residential segregation indirectly by maintaining or increasing income inequality and directly by further stratifying real estate values.*

### Demographic Factors and Segregation

*Migration.*—Cross-sectional country comparisons and the historical experience of the currently industrialized countries shows that industrialization and urbanization are often thought to move in tandem (Gilles et al. 1992). However, this relation does not hold in currently less-developed countries, including Brazil (Firebaugh 1979; Merrick and Graham 1979). Third World cities grow largely through natural increase, which far outpaces the growth of cities in industrialized countries during their peak growth periods, and migration is often more related to adverse conditions in rural areas than to the attraction of urban areas, as exemplified by the rapid growth of cities in the least-industrialized regions of Brazil (Martine and Camargo 1984). Also, social networks between an urban area and other areas, which may have started when an urban area was economically attractive to migrants, may continue to reinforce migration despite withering economic opportunities. Thus, demographic factors such as population size and migration should be analyzed separately from industrialization.

Population growth from migration may increase segregation by overburdening the formal housing market and increasing competition for cen-

tral land and housing. Also, since migrants tend to enter the urban labor market without financial capital and as low-wage and often informal-sector workers, they are often limited to poor housing. Their strong social networks further channel them into particular neighborhoods that are often populated with many other poor migrants, thus increasing economic residential segregation (Perlman 1976; Leeds 1974). Population growth forces new neighborhoods to be formed, whether these be shantytowns or middle-class settlements, and because these new neighborhoods are especially homogeneous in earlier stages, greater segregation is likely to accompany such growth. Finally, human ecologists have noted that population growth intensifies demand for central land and increases spatial differentiation, thus increasing segregation (Shevky and Bell 1955; Guest 1984). Thus

**HYPOTHESIS 5.**—*Higher migration rates or faster growth of metropolitan areas leads to greater segregation.*

**Population size.**—The major difference between urban areas today and in the past is size—contemporary large urban areas are far larger and more numerous than in the past. Currently, many less-developed countries feature giant or mega-cities (Angiotti 1993). Brazil currently has two mega-cities: São Paulo, with more than 12 million inhabitants in 1980, and Rio de Janeiro, with about 9 million residents.

Human ecologists have stressed the role of population size in increasing segregation (Schnore 1958). Population size allows for the development of greater social differentiation as well as greater differentiation of neighborhoods (Hawley 1950; Choldin 1984; White 1986a). Also, larger urban areas tend to have greater commuting distances and times, increasing demand for neighborhoods with better access to transportation routes and consumer and labor markets. This should lead to greater disparities in land values in larger areas, thus increasing segregation. Also, higher values overall would likely increase the chances of separate informal or illegal housing markets. This is clear in Brazil, as its rapid urban population growth in the 1970s often outpaced the housing supply. Thus

**HYPOTHESIS 6.**—*Larger urban areas tend to have greater residential segregation.*

### The Special Case of Brasilia

Any analysis of segregation in Brazilian metropolitan areas should note the special case of Brasilia, which replaced Rio de Janeiro as the new national capital in 1960 and was to represent Brazil as a modern nation. Brasilia was designed, built, and administered by the state in the interest of, among other things, subverting the pernicious residential segregation and social class divisions that characterized typical Brazilian cities (Hol-

sten 1988). Brasilia's creators had a vision of a residentially egalitarian city with persons of distinct social classes living together in standardized apartment buildings, which themselves were intended to negate social divisions (Holsten 1988). However, its transformation to a free real estate market and the growth of the Brasilia metropolitan area beyond the limits of the planned city (Plano Piloto) has led to very apparent class segregation between primarily middle-class government bureaucrats, who can afford to live in Brasilia city, and the low-skilled manual workers, who must live in the satellite cities (Campos 1991).<sup>3</sup> Thus, the spatial rigidity imposed by the highly ordered housing zones may have instead exaggerated Brasilia's spatial segregation beyond that of older Brazilian cities. Thus, the final hypothesis:

**HYPOTHESIS 7.**—*Brasilia has significantly higher segregation than other Brazilian metropolitan areas.*

## DATA AND VARIABLES

### Data

Data are from the microlevel data files of the 1980 census of Brazil. All information on segregation is computed from the full 25% sample. Indexes of segregation were computed by the Instituto Brasileiro de Geografia e Estatística (IBGE), under close supervision by the author. They are based on census tracts that have between 250 and 300 households, or an average of 1,150 persons.<sup>4</sup> Also, census tracts were designed in or prior to 1960 and new subdivisions are made according to physical criteria, thus diminishing the effects that variation in census tract design would have on measured social segregation across the metropolitan areas. The uniformity in the size of census tract parcels and fairly uniform criteria for their design make intranational comparisons viable. The units of analysis are the 40 Brazilian urban areas in 1980, as specified in an IBGE publication (Vetter 1988), with populations greater than 200,000.

### Measures of Segregation

To measure segregation, I examine the extent of *evenness* in the distribution of household income groups across metropolitan areas. The concept

<sup>3</sup> Although the initial plan sought the integration of members of diverse classes, there was a clear differentiation of housing types between neighborhoods. Thus, when market forces were allowed to take over, sharp differences in housing values emerged among neighborhoods.

<sup>4</sup> By contrast, U.S. urban census tracts average about 5,000 persons. The larger size tends to reduce overall segregation values, as larger tracts are likely to be more heterogeneous.

of evenness refers to the extent to which social groups are differentially distributed across a metropolitan area. It is particularly suitable for capturing the amount of segregation found among the mosaics of households and neighborhoods that characterize the landscapes of Brazilian metropolitan areas. Much research on Latin American patterns of segregation have focused on *centralization* instead, largely because it describes an urban form derived from colonial (and even precolonial) times that is relatively easy to observe and, if measured, requires data for only two areas: the central city and the periphery. However, centralization fixes an a priori urban form that is both overly simplistic and often inappropriate for describing Brazil's spatially complex metropolitan areas.

The traditionally employed measure of evenness, and of residential segregation in general, is the widely used and intuitively interpretable index of dissimilarity ( $D$ ) (Massey and Denton 1988; White 1986b). However,  $D$  is limited to the comparison of two groups. Because residential segregation by household income involves at least several categories, several pairwise comparisons must be made with  $D$ . To obtain a single citywide measure of segregation by social class, individual  $D$ s must be aggregated as analysts have done in the past (Farley 1977; Denton and Massey 1988), or an index that can appropriately handle polytomous data must be chosen. Because  $D$  is sensitive to random fluctuations in the distribution of persons across census tracts, its aggregation to a single measure may be inaccurate. Several authors have suggested a rarely used entropy or information-based index known as the entropy measure of segregation or, simply,  $H$  (James and Taueber 1985; White 1986b; Massey and Denton 1988). Specifically,  $H$  measures departure from evenness by taking the weighted mean deviation of every census tract's entropy ( $E$ ) from the entropy of the entire metropolitan area.

Metropolitan area-wide entropy, a summary measure of (income) composition, is computed as

$$E = -1 * \sum (P_k \log P_k),$$

and a census tract's entropy ( $E_i$ ) is similarly

$$E_i = \sum (p_{ik} \log p_{ik}),$$

and finally, the entropy index of residential segregation ( $H$ ) is

$$H = \sum (p_i(E - E_i)/EP) * 100,$$

where  $p_i$  and  $p_{ik}$  refer, respectively, to the census tract ( $i$ ) population and the proportion of the population of each income group ( $k$ ) in each census tract. The variables  $P$  and  $P_k$  similarly refer, respectively, to the total population of a metropolitan area and the proportion of the population of each income group among the total metropolitan area population.



TABLE 1

DISSIMILARITY INDEXES AMONG HOUSEHOLD INCOME GROUPS BY NUMBER OF  
MINIMUM WAGES IN FOUR METROPOLITAN AREAS OF BRAZIL: 1980

Metropolitan Area and Income Group	% of Total Population	% Nonwhite	20+	10-19.99	5-9.99	3-4.99	2-2.99	1-1.99
São Paulo:								
20+ .....	7.3	2.9	..					
10-19.99 .....	14.6	10.7	32.0	.				
5-9.99 .....	28.1	22.2	32.9	20.7	.			
3-4.99 .....	23.8	31.7	34.1	24.7	19.1	..		
2-2.99 .....	13.2	37.6	39.5	35.6	30.6	21.1	..	
1-1.99 .....	10.0	41.1	54.1	54.2	51.4	44.5	30.3	...
Less than 1 .....	2.9	36.0	73.7	73.9	73.2	69.5	60.1	41.4
Rio de Janeiro:								
20+ .. .....	6.7	5.1	...					
10-19.99 ....	11.1	15.1	24.5	...				
5-9.99 .....	21.2	31.0	27.8	18.4	...			
3-4.99 .....	22.2	42.6	33.2	25.6	19.7	..		
2-2.99 .....	16.1	48.3	46.8	42.3	36.5	26.0	...	
1-1.99 .....	16.9	52.2	67.4	65.4	61.5	53.4	36.1	...
Less than 1 .....	5.8	52.8	83.4	82.7	81.0	76.7	65.3	42.0
Salvador:								
20+ .....	6.1	28.2	...					
10-19.99 .....	10.4	48.0	22.5	...				
5-9.99 .....	17.6	68.0	26.2	19.4	...			
3-4.99 ....	18.4	79.7	31.2	27.6	21.3	...		
2-2.99 .....	16.0	84.9	44.8	44.5	38.5	27.6	..	
1-1.99 .....	19.7	87.6	65.7	67.4	63.4	55.7	38.0	..
Less than 1 .....	11.8	89.1	81.4	83.6	82.0	77.2	65.6	42.5
Brasília:								
20+ .....	10.3	10.7	..					
10-19.99 .....	12.8	24.5	27.8	...				
5-9.99 ... ..	19.3	42.8	31.6	20.2	...			
3-4.99 .....	19.8	55.1	38.0	28.6	20.8	...		
2-2.99 .....	15.7	60.1	50.2	46.9	40.3	29.2	..	
1-1.99 .....	17.2	62.9	68.7	70.9	67.5	59.6	38.8	...
Less than 1 .....	4.8	64.9	83.0	86.2	85.7	82.9	72.7	48.5

In addition to  $H$ , I present three other indexes of overall segregation based on  $D$ . The main substantive difference between  $H$  and the measures based on  $D$  is that  $H$  measures the extent to which the income composition of census tracts deviates from the citywide composition, while the  $D$  measures are based on the extent to which paired groups deviate from each other in their distribution across census tracts. The first two indexes are aggregates of all pairs (as presented in table 1) but are weighted in different ways. "Standardized mean  $D$ " is weighted by

a standard or average population. I choose the metropolitan area of Juiz de Fora as the standard population since it very closely matches the mean income distribution of the 40 metropolitan areas. Thus, standardized mean  $D$  is the weighted mean dissimilarity of a metropolitan area if it had the income distribution of Juiz de Fora. The second index, "weighted mean  $D$ ," is weighted by the own income composition of each metropolitan area. The third index, " $D$  between extremes," is simply the dissimilarity index between the highest and lowest income groups.

Weighted mean  $D$ , because it mathematically incorporates income composition, captures a different conception of segregation than the other measures. Specifically, it reflects the idea that, even though the highest and lowest income groups may be similarly segregated from each other in two metropolitan areas, the fact that these groups represent a higher proportion of the population in city A than in city B should make for greater segregation in city A, assuming all other groups are equally segregated. Standardized mean  $D$ , though, would yield the same segregation levels between the two cities, and  $H$  would yield similar levels.

### Independent Variables

To estimate relative levels of industrialization, I use the percentage of the total labor force employed in manufacturing. The Gini index of income inequality is estimated from the midpoint values of the lowest six ranked income intervals as shown in table 1 and from assigning a value of 25 for the highest, open-ended interval. Population size is logged to pull in large values, especially those of the two giant metropolitan areas. Percentage migrant refers to the total resident migrant population divided by the total population.

## FINDINGS

### Segregation in Four Metropolitan Areas

This section describes levels of segregation between paired household income groups in São Paulo, Rio de Janeiro, Salvador, and Brasília. The use of dissimilarity indexes that pair all combinations of the income groups provides a sense of the nature and extent of segregation in various urban contexts. Table 1 shows that residential segregation follows the expected pattern of greater spatial distance with greater income differences.<sup>5</sup>

<sup>5</sup> Segregation indexes were also computed for broad occupational groups. Bivariate and multivariate results based on the aggregate indexes ( $D$  and  $H$ ) for occupational groups were similar to those based on income. The greatest exception to the pattern expected from segregation by income was in the examination of the paired dissimilarity scores. Segregation of white-collar workers from personal-service workers was lower than from sales, skilled manual, and transport workers, even though the last three

Using a rule of thumb in which values of 70–100 indicate extreme segregation (Massey and Denton 1987), extreme segregation is found between households earning less than one minimum wage and (1) *five* income categories in Brasília,<sup>6</sup> (2) *four* categories in Rio de Janeiro and Salvador, and (3) *three* categories in São Paulo. The second column shows that households earning less than one minimum wage represent only 2.9% of São Paulo's population, and aside from this very small group, extreme segregation among income groups is not found in that metropolitan area. By contrast, the proportion of Salvador's population earning less than one minimum wage is fully 11.8% and is extremely segregated from most of Salvador's remaining population—those earning over three minimum salaries. In addition to the lowest income category, extreme or nearly extreme levels are found in the next poorest, but substantially larger, income category (one to two minimum wages) in Rio de Janeiro, Salvador, and Brasília. In Brasília, the two lowest income groups, which together represent 22% of Brasília's population, are extremely or almost extremely segregated (values of 67.5 and higher) from the 43% of the population that earns at least five minimum wages.<sup>7</sup> Thus, segregation of the middle class from the poor is moderate in highly industrialized São Paulo, substantially greater in Rio de Janeiro and Salvador, and clearly greatest in Brasília.

Further analysis of segregation among proximate groups at high income levels reveals a different pattern. Among the three highest earning groups, São Paulo has the greatest segregation at this level. Thus, segregation between the upper and lower middle classes, in rough terms, is greatest in São Paulo and lowest in Rio de Janeiro, although segregation between the poor and the middle class is greater in Rio de Janeiro than in São Paulo.

---

groups have higher incomes, on average, than personal-service workers. This may reflect the fact that personal-service workers include domestic workers, custodians, and security guards who often live on the premises of their employers.

<sup>6</sup> Numbers of minimum wages is a convenient way to collect and represent data in Brazil because of its high rates of inflation. In 1980, one minimum wage equaled about 75 U.S. dollars.

<sup>7</sup> The relatively low levels of segregation in São Paulo are reflected in *D* between those earning 10–20 minimum wages and those earning 1–2 minimum wages. This is an important comparison of middle-class and poor groups in all four areas because these income groups represent substantial portions of the population in all four areas. São Paulo, e.g., has a *D* in this case of 54.1, compared to 67.4 for Rio de Janeiro. Differences in segregation between the poor and middle classes in São Paulo and Rio de Janeiro are even greater if we consider that the poor groups represent a much larger portion of the population in Rio de Janeiro, so that a much larger portion of the population in Rio would have to exchange census tracts with the middle class to achieve evenness at this level.

Class and race are closely correlated in Brazil so that class issues, like segregation, often become racialized in many Brazilian cities. The wealthiest groups are often almost all white and the poor are disproportionately nonwhite, although racial composition and thus racial composition by class may vary widely across cities. The third column of table 1 shows the racial composition of income groups in the four metropolitan areas. In Salvador, where roughly 80% of the population identifies itself as nonwhite, nonwhites constitute 80%–90% of the lowest four income groups but less than 30% of the highest earning group. In São Paulo, nonwhites are a large minority of about 30%–40% among the poor, but they are virtually absent (only 2.9%) among the highest income group. Thus some racial segregation in Brazil is assured by the fact that nonwhites are disproportionately in lower socioeconomic groups, although moderate racial segregation also occurs among members of the same income category (Telles 1992).

#### Levels of Overall Segregation in 40 Metropolitan Areas

While the previous analysis has focused on segregation between paired income groups within metropolitan areas, the remaining analysis examines overall segregation in the 40 metropolitan areas. Table 2 shows levels of segregation as measured by the four indexes and levels of income inequality for the 40 largest metropolitan areas by region and then by order of population size. Levels of segregation tend to be highest in the least developed Northeast region, although the relation is not very strong. Consistent with greater spatial inequality, levels of income inequality among Northeast urban areas are invariably higher than those of urban areas in other regions, with the single exception of nearby Belem. In fact, the lowest levels of income inequality tend to be in São Paulo, the most industrialized state of Brazil. Finally, table 2 shows that segregation tends to decrease with population size.

Values of  $H$  are smaller than those of  $D$ , which tends to be true in general (White 1986a). Also, income segregation values using  $H$  are likely to be especially small because income groups are often continuous, resulting in some *mixing* across the boundaries of the categories.

Brasilia has the highest levels of income segregation under  $H$  (20.1) and standardized mean  $D$  (50.7), despite having about average levels of income inequality (Gini of 48.3). Thus, Brasilia's relatively high segregation seems to result from high levels of residential segregation between paired income groups (table 1) rather than from overall income inequality. Teresina scores the highest on the weighted dissimilarity index (54.5), apparently because it has the highest level of income inequality (Gini of 56.1) and income inequality is directly incorporated into this segregation

TABLE 2  
INDEXES OF SEGREGATION AND GINI INDEX OF INCOME INEQUALITY FOR THE 40 LARGEST METROPOLITAN AREAS OF BRAZIL: 1980

REGION/STATE AND METROPOLITAN AREA	ENTROPY INDEX (H)	DISSIMILARITY (D)			GINI INDEX
		Standardized Mean*	Weighted Mean	Between Extremes†	
Northeast region:					
Recife .....	15.3	49.7	53.2	84.1	54.6
Salvador .....	15.6	47.5	48.4	81.4	51.3
Fortaleza .....	13.3	46.9	50.4	80.2	54.4
João Pessoa ..	14.9	50.0	53.9	85.7	55.3
Natal .....	14.8	48.9	51.5	80.5	53.4
Teresina .....	14.4	49.5	54.5	82.0	56.1
Maceio .....	13.0	47.3	51.2	81.6	55.0
Aracaju .....	12.2	44.5	46.8	78.0	52.9
São Luis .....	10.2	43.0	46.8	74.9	54.3
Feira de Santana .....	8.5	39.3	41.8	71.5	51.5
Campina Grande .....	10.9	44.7	49.4	76.4	55.8
Itabuna .....	12.3	45.5	48.7	80.9	54.1
Minas Gerais—Rio de Janeiro—Espírito Santo states:					
Rio de Janeiro .....	14.8	46.7	45.4	83.4	47.3
Belo Horizonte .....	15.2	46.3	45.4	81.1	48.0
Vitoria .....	14.5	47.4	47.6	82.6	49.1
Barra Mansa—Volta Redonda .....	11.6	42.9	41.5	76.5	44.6
Juiz de Fora .....	11.0	41.7	41.7	77.0	48.6
Ipatinga .....	14.0	45.3	44.7	75.8	46.7
Uberlandia .....	9.5	39.4	38.7	68.0	46.5
Campos .....	9.6	40.4	41.5	72.6	49.9

São Paulo state:

São Paulo .....	12.7	41.7	39.5	73.7	42.3
Santos .....	12.3	43.5	41.6	74.9	42.1
Campinas .....	12.7	42.2	40.3	71.0	43.1
São José dos Campos .....	10.7	39.1	37.2	69.4	43.5
Sorocaba .....	8.1	36.4	34.3	66.2	42.1
Ribeirão Preto .....	11.9	41.5	39.3	76.0	44.0
Jundiaí .....	9.2	38.9	37.0	67.0	40.5
South region.					
Porto Alegre .....	13.1	44.7	43.0	77.4	45.1
Curitiba .....	12.2	43.3	42.0	73.3	46.2
Pelotas-Rio Grande .....	11.1	44.7	44.5	77.3	46.9
Florianópolis .....	12.3	41.8	41.0	74.5	47.8
Londrina .....	13.9	46.0	45.5	77.9	48.0
Joinville .....	8.9	40.3	38.2	64.6	41.3
Caxias do Sul .....	9.8	40.4	38.4	70.4	41.6
North and central west (frontier) regions.					
Brasília .....	20.1	50.7	49.4	83.0	48.3
Belém .....	11.4	43.4	45.1	78.3	51.6
Goiânia .....	11.8	42.2	42.7	75.6	49.9
Manaus .....	10.3	40.4	39.7	66.7	46.9
Campo Grande .....	11.2	40.2	39.7	67.9	48.5
Cuiabá .....	11.5	41.0	41.4	69.3	50.2
Mean .....	12.1	43.6	43.9	75.5	48.5

\* Standard metropolitan area is Juiz de Fora.

† Extreme income groups: households earning 20 or more minimum wages and households earning less than 1 minimum wage.

TABLE 3

## ZERO-ORDER CORRELATIONS AMONG RESIDENTIAL SEGREGATION MEASURES

Segregation Measures	1	2	3	4
1. Entropy index ( <i>H</i> )	1.000			
2. Standardized mean <i>D</i>	.885	1.000		
3. Weighted mean <i>D</i>	.714	.934	1.000	
4. <i>D</i> between extremes	.828	.917	.865	1.000

NOTE.—*N* = 40. Values over .316 significant at the *P* < .05 level

measure. In terms of *D* between extremes, only Recife (84.1) and Rio de Janeiro (83.4) have higher levels than Brasília (83.0).

Brasília, Recife, Salvador, João Pessoa, and Teresina score among the 10 most segregated metropolitan areas on all four measures. Aside from Brasília, the other four areas are all in the underdeveloped Northeast region and are among the largest 50%. Places having among the 10 lowest levels of segregation on all four measures are São José dos Campos, Sorocaba, Jundiaí, Uberlândia, Joinville, and Caxias do Sul, all of which are relatively small and in the more industrialized regions. These results suggest that size is positively associated and industrialization negatively associated with segregation.

For purposes of the multivariate analysis, I chose *H* to represent segregation, a decision consistent with previous prescriptions for its use (White 1986b; Massey and Denton 1988). The use of *H* rather than an aggregate *D* measure seems to be further justified by the fact that all four measures are empirically correlated, so that the choice of one over the other is not likely to greatly affect the findings. As table 3 shows, the four measures of segregation are highly correlated among themselves. Measure *H* is very consistent with standardized mean *D* ( $r = .885$ ), reflecting their conceptual similarity. Measure *H* correlates the least with weighted mean *D* ( $r = .714$ ) because local income inequality is mathematically built into the latter measure, an undesirable property in an analysis where income inequality is a covariate. Finally, *D* between extremes, although it strongly correlates with the aggregate measures, is selective of only the wealthiest and poorest sectors of the population.

### Explaining Patterns of Segregation

To test how both industrialization and urbanization affect segregation, I regress *H* on the independent variables through four models. *H* is transformed into logits because of its limited range. The units of analysis are 39 of the 40 metropolitan areas. I exclude Brasília from the analysis

# Socioeconomic Segregation in Brazil

TABLE 4

OLS REGRESSION RESULTS PREDICTING ENTROPY MEASURE  
OF RESIDENTIAL SEGREGATION ( $H$ )

INDEPENDENT VARIABLES	MODEL				MEAN	SD
	1	2	3	4		
% in manufacturing .....	-.006** (.002) [-.393]	-.003 (.004) [-.212]	-.007*** (.004) [-.459]	-.003 (.003) [-.180]	20.9	12.5
Gini index of inequality . .		.010 (.010) [.238]		.016** (.008) [.390]	48.5	4.6
Population (logged) .....			.119*** (.025) [.614]	.127*** (.024) [.653]	13.2	1.0
% migrant .....			.004 (.003) [.180]	.006* (.003) [.252]	23.3	8.1
Intercept . . . . .	-1.871	-2.417	-3.683	-4.556		
$R^2$ .....	.154	.178	.491	.551		
Adjusted $R^2$ .....	.132	.132	.448	.498		

NOTE —  $N = 39$  Dependent variable transformed into logits, SEs are in parentheses, and standardized coefficients are in brackets

\*  $P < .10$   
 \*\*  $P < .05$   
 \*\*\*  $P < .01$

because of its extreme segregation and statistical outlier status, as regression diagnostics indicated.

Regression results are shown in table 4. Model 1 shows that residential segregation decreases with industrialization, and model 3 shows an even stronger negative relationship when population size and migration are controlled. However, the addition of the income inequality variable in the simpler model (model 2) results in the loss of predictive power for industrialization, although inequality is not statistically significant in either. In the full model (model 4), industrialization is not significant, while inequality and both demographic variables are. Inequality is positively related to segregation, demonstrating that segregation directly reflects the extent of income inequality in an area of similar size and migration. This also reflects the modest correlation between industrialization and inequality, in which half of the variance between the two variables overlaps ( $r = -.76$ , in the appendix), indicating that greater industrialization tends to reduce inequality among metropolitan areas. Model 4 also shows that larger urban areas and those with larger proportions of migrants tend to be more segregated. The inclusion of demographic variables (models 3



TABLE 5

OLS REGRESSION RESULTS PREDICTING INCOME INEQUALITY

INDEPENDENT VARIABLES	MODEL	
	1	2
% in manufacturing .....	-.283*** (.040) [-.760]	-.266*** (.041) [-.715]
Population (logged) .....		-.466 (.515) [-.186]
% migrant .....		-.108 (.066) [-.101]
Intercept .....	54.397	62.695
$R^2$ .....	.578	.609
Adjusted $R^2$ .....	.566	.576

NOTE.— $N = 39$  Dependent variable transformed into logits; SEs in parentheses, and standardized coefficients in brackets.

\*  $P < .10$ .

\*\*  $P < .05$ .

\*\*\*  $P < .01$ .

and 4) more than tripled the explanatory power ( $R^2$ ) of the simpler models (1 and 2), so that about half of the variation in segregation was explained.<sup>8</sup>

The direct effect of industrialization on inequality is shown in table 5. Inequality is clearly reduced when greater proportions of the labor force are employed in manufacturing industries, and over half of the varia-

<sup>8</sup> Several other variables, such as those representing the housing market, were considered but later dropped because they tended to be less theoretically defensible than the included variables or had empirically insignificant effects. Housing turnover, which has been found to be related to class segregation (O'Loughlin 1983), is highly correlated with percentage migrant ( $r = .81$ ), and homeownership has virtually no effect in any model ( $t < 1$ ). Also, age has been shown to be an important predictor of segregation in the United States. However, a dummy variable for age, operationalized as whether the *município* population reached 100,000 by 1940 (the year of Brazil's first reliable census), had no independent effect on segregation ( $t < 1$ ). Finally, indexes of racial segregation were unrelated to class segregation ( $t < 1$ ), a finding similar to that for the United States (Elgie and Clark 1981). I was also concerned that high density in urban areas had distinct implications for segregation because the physical area of a census tract is smaller than in lower-density areas. The percentage in apartments is an indicator of density and varies from 0.4 in Teresina to 34.3 in Santos. However, a separate regression analysis that excluded the four urban areas with more than 15% of their population living in apartments (more than 1 SD above the mean) yielded results similar to the findings for the 39 areas.

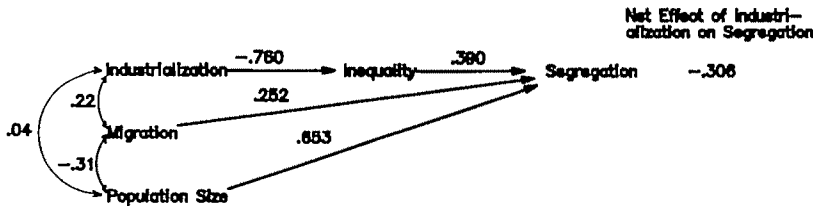


FIG. 2.—Path diagram showing effects of structural variables on segregation

tion in inequality is explained (model 1). Finally, the inclusion of demographic variables had virtually no effect on income inequality (model 2).

Thus, the results support hypothesis 1 that industrialization reduces segregation but only by reducing income inequality. Larger populations and a higher share of migrants also mean more segregation, supporting hypotheses 5 and 6. Furthermore, hypothesis 1 is supported only when the demographic effects are controlled. The relative effect of the variables is approximated with standardized coefficients shown in brackets in tables 4 and 5, and the final results are illustrated in the path diagram in figure 2. Population size clearly has the strongest direct effect on segregation (.653), followed by inequality (.390) and migration (.252). Industrialization has a relatively strong negative effect on inequality ( $b = -.760$ ). Thus, the indirect effect of industrialization on segregation in this model is  $-.306$  ( $.390 \times -.760$ ), an effect that is substantially less than that of population size.

## CONCLUSIONS

Industrialization is associated with lower segregation across Brazilian metropolitan areas, but its effect occurs only indirectly, via an area's income inequality. The cross-sectional finding of a strong negative relation between industrialization and income inequality thus supports neo-classical economic theory. It suggests that workers' wages in the aggregate are greater relative to the middle class in Brazil in places where a larger share of the labor force is incorporated into industrial jobs. Thus, more industrialized areas are likely to have lower income inequality than less industrialized places at any single point in time.

Industrialization has no discernible direct effects on segregation, as would be expected from theories claiming that industrial location geographically clusters workers (Schnore 1965; Fales and Moses 1972; Logan and Molotch 1987) or that industrial profits are invested into local real estate markets, thus heightening spatial stratification (Harvey 1978; Lo-

gan and Molotch 1987; Kowarick and Campanario 1988; Geisse and Sabatini 1988; Ribeiro 1992). Because proponents of these theories often draw their evidence from studies of a single case, they may be unable to discern the overriding effects of changing income inequality on segregation from these other effects.

More important than industrialization, urbanization explains much more of the variation in segregation than industrialization and inequality. In particular, population size is the best predictor of an area's segregation, making isolation of the poor and working classes most likely to be greatest in the largest cities. The importance of demographic variables in explaining segregation has been ignored or understated in many political economy discussions of the determinants of segregation, which focus on variables like industrialization and inequality. Urbanization is often assumed to move in tandem with industrialization, but the processes are often independent of each other in the currently less industrialized countries, and their effects on segregation may occur in opposite directions, as the case of Brazil shows. Industrialization decreases segregation, while urbanization increases it. Furthermore, the industrialization-inequality-segregation effect emerges only when population size and migration are controlled.

Given the strong relation between segregation and social well-being, slowed growth among the largest Brazilian metropolitan areas and greater growth among smaller ones in recent years therefore gives reason for some optimism. Debates over the effects of the large size of cities in the Third World focus on economic and health issues but tend to overlook social issues. The emerging "mega-cities" of the Third World are thus especially prone to a high degree of socioeconomic segregation.

On the other hand, this pattern may not endure into the future. The transition from large-scale production to highly competitive, small-scale and flexible industrial production may mean greater income inequality, as the case of some global cities suggests (Sassen 1990), and thus greater segregation. Given substantial inequality in even the industrialized areas of countries like Brazil, new forms of industrialization may not affect inequality and may actually reduce the inequality and segregation of places that currently have little or no industrialization. Brazil is only beginning to enter this stage of industrialization, and thus how it will affect urban development, inequality, and segregation is still far from clear.

Income segregation in Brazil implies racial segregation because the two are correlated, although racial segregation is also created by moderate levels of segregation between similar socioeconomic groups (Telles 1992). Also, spatial factors may help explain why class identities are stronger than racial identities. In places like Rio de Janeiro and Brasilia, segrega-

tion between the poor and the middle class is extreme, while racial segregation is not. Furthermore, while nonwhites are more likely than whites to be poor, white-nonwhite segregation among the poor is especially low (Telles 1992).

The fact that metropolitan areas in the underdeveloped Northeast region have the highest levels of income inequality and segregation and that residents in that region are predominately nonwhite means that, not only are nonwhites more likely to suffer from poverty than whites, but they are also more disadvantaged by virtue of living in more unequal and segregated places. The Northeast's underdevelopment and the predominance of Afro-Brazilians in the Northeast are dual but related outcomes of the region's underdevelopment, historically rooted in a system of large landholdings for the development of a single export crop—usually sugar—which in turn depended on the brute labor of nonwhite workers.

Most of the variation in segregation among Brazilian metropolitan areas is explained in this analysis by industrialization and urbanization, although much of it may be further explained by the local management decisions and physical idiosyncrasies of particular cities and regions. Although housing and labor policies tend to be uniform throughout Brazil, differences in the enforcement of these may also help explain urban variation in segregation. Physical peculiarities like the mountainous terrain of Rio de Janeiro may help shape segregation, although Rio's segregation fits the modeled pattern based on all metropolitan areas.

The case of Brasilia offers an important lesson in urban planning. Its particularly high level of segregation originates in its unique development, which has overcome the constraints of industrialization and urbanization that predict segregation in other Brazilian urban areas. Segregation in Brasilia has been shaped by a highly ordered urban design, followed by conversion to a free real estate market and unanticipated population growth beyond the planned city limits. Ironically, Brasilia's planners had sought to build a new capitol free of the enormous class distances found in other Brazilian cities.

The direction of causality between industrialization, inequality, and segregation may also flow in reverse, although such effects are not likely to be observed at the metropolitan area level but at national or regional levels. Brazil's large income inequality, as well as that of other Latin American countries, may itself be an obstacle to development. According to some economists, the strong pressure in highly unequal societies for populist solutions has led to the avoidance of the currency devaluations that were necessary for Taiwan and Korea to attain export competitiveness and reach their current high levels of development (Sachs 1989; Mahon 1992). Similarly, high levels of segregation may impede develop-

ment as poor access to labor markets for many workers reduces their employability and lowers industrial efficiency. Also, inadequate access to good schools and other public services, as well as the psychological effect that segregation has on children and adolescents, limits the formation of human capital that is necessary for development.

Finally, given the complexity of residential form and segregation in the metropolitan areas of Third World countries, this study has demonstrated the need for indexes based on census tract data to measure levels of segregation. Such work could be extended to measure the social impact of segregation on the urban populations of these countries, particularly on the poor and on ethnic groups. Furthermore, because of the sensitivity of these measures to the size and design of census tract parcels, efforts should be made to create comparable tracts across countries. This would allow cross-national comparisons of cities, work that is currently hindered because of a general lack of uniformity at this level of aggregation. The choice of indexes has been shown to be important for accurately measuring segregation, as assumptions and limitations in their construction are not readily apparent.

## APPENDIX

TABLE A1  
ZERO-ORDER CORRELATIONS AMONG ALL INDEPENDENT VARIABLES  
USED IN THE REGRESSION ANALYSIS

Variable	1	2	3	4	5
1. Entropy index ( $H$ ) ....	1.00				
2. % in manufacturing .....	-.39	1.00			
3. Gini index of inequality .....	.40	-.76	1.00		
4. Population (logged) .....	.54	.04	-.07	1.00	
5. % migrant ..	-.11	.22	-.31	-.31	1.00

NOTE.— $N = 39$ . Values greater than .316 significant at the  $P < .05$  level.

## REFERENCES

- Angiotti, Thomas. 1993. *Metropolis 2000: Planning, Poverty and Politics*. New York: Routledge.
- Bava, Silvio Caccia. 1994. "Neighborhood Movements and Trade Unions: The São Bernardo Experience." In *Social Struggles and the City: The Case of São Paulo*, edited by Lucio Kowarick. New York: Monthly Review Press.
- Bonelli, Regis, and Guilherme Luís Sedlacek. 1989. "Distribuição de Renda: Evolução no Último Século." In *Mercado de Trabalho e Distribuição de Renda: Uma Coletânea*, edited by Guilherme Luís Sedlacek and Ricardo Paes de Barros. Rio de Janeiro: Instituto de Planejamento Econômico e Social.
- Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela Kato Klebanov, and Naomi Sealander. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99 (2): 353-95.
- Caldeira, Teresa Pires do Rio. 1992. "City of Walls: Crime, Segregation and Citizenship in São Paulo." Ph.D. dissertation. University of California, Berkeley, Department of Anthropology.
- Campos, Neio. 1991. "A Segregação Planejada." In *A Conquista da Cidade: Movimentos Populares em Brasília*, edited by Aldo Paviani. Brasília: Editora Universidade de Brasília.
- Chenery, Hollis B. 1980. "Poverty and Progress—Choices for the Developing World." *Finance and Development* 17 (2): 12-16.
- Choldin, Harvey M. 1984. "Subcommunities: Neighborhoods and Suburbs in Comparative Perspective." Chap. 6 in *Sociological Human Ecology: Contemporary Issues and Applications*, edited by Michael Micklin and Harvey Choldin. Boulder, Colo.: Westview.
- Denton, Nancy A., and Douglas S. Massey. 1988. "Residential Segregation of Blacks, Hispanics and Asians by Socioeconomic Status and Generation." *Social Science Quarterly* 69:797-817.
- Elgie, Robert A., and Alex Rees Clark. 1981. "Social Class Segregation in Southern Metropolitan Areas." *Urban Affairs Quarterly* 16 (3): 299-316.
- Engels, Friedrich. (1845) 1973. *The Condition of the Working Class in England in 1844*. Moscow: Progress Publishers.
- Evans, Peter. 1979. *Dependent Development: The Alliance of Multinational, State and Local Capital in Brazil*. Princeton, N.J.: Princeton University Press.
- Fales, R. I., and L. N. Moses. 1972. "Land Use Theory and the Spatial Structure of the Nineteenth Century City." *Regional Science Association Papers* 28:49-80.
- Farley, Reynolds. 1977. "Residential Segregation in Urbanized Areas of the United States in 1970: An Analysis of Social Class and Racial Differences." *Demography* 14:497-518.
- Firebaugh, Glenn. 1979. "Structural Determinants of Urbanization in Asia and Latin America." *American Sociological Review* 44:195-215.
- Folha de São Paulo. 1992. "Arrastão Apavora a Classe Média." (October 25).
- Geisse, Guillermo, and Francisco Sabatini. 1988. "Latin American Cities and Their Poor." Chap. 11 in *The Metropolis Era: A World of Giant Cities*, edited by Mattei Dogan and John D. Kasarda. Newbury Park, Calif.: Sage.
- Gilles, Malcolm, Dwight H. Perkins, Michael Roemer, and Donald R. Snodgrass. 1992. *Economics of Development*, 3d ed. New York: Norton.
- Guest, Avery. 1984. "The City." Chap. 7 in *Sociological Human Ecology: Contemporary Issues and Applications*, edited by Michael Micklin and Harvey Choldin. Boulder, Colo.: Westview.

- Harvey, David. 1985. "The Urban Process under Capitalism: A Framework for Analysis." Chap. 1 in *The Urbanisation of Capital*, by David Harvey. Baltimore: Johns Hopkins University Press.
- Hawley, Amos. 1950. *Human Ecology: A Theory of Community Structure*. New York: Ronald
- Holsten, James. 1988. *The Modernist City*. Chicago: University of Chicago Press.
- James, David R., and Karl E. Taueber. 1985. "Measures of Segregation." *Sociological Methodology* 15:1-32.
- Kowarick, Lucio, and Clara Ant. 1988. "Cem Anos de Promiscuidade: O Cortiço na Cidade de São Paulo." Chap. 2 in *As Lutas Sociais e a Cidade*, edited by Lucio Kowarick. São Paulo: Paz e Terra.
- Kowarick, Lucio, and Milton A. Campanário. 1988. "São Paulo: Metrópole do Subdesenvolvimento Industrializado." In *As Lutas Sociais e a Cidade*, edited by Lucio Kowarick. São Paulo: Paz e Terra.
- Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45:1-28
- Leeds, Anthony. 1974. "Housing Settlement Types, Arrangements for Living, Proletarianization, and the Social Structure of the City." In *Latin American Urban Research: Anthropological Perspectives on Latin American Urbanization*, edited by Wayne A. Cornelius and Felicity M. Trueblood. Beverly Hills, Calif.: Sage.
- Logan, John R., and Harvey L. Molotch. 1987. *Urban Fortunes: The Political Economy of Place*. Berkeley and Los Angeles: University of California Press.
- Mahon, James E. 1992. "Was Latin America Too Rich to Prosper? Structural and Political Obstacles to Export-Led Industrial Growth." *Journal of Development Studies* 28:241-63.
- Martine, George, and Liscio Camargo. 1984. "Crescimento e Distribuição da População Brasileira: Tendências Recentes." *Revista Brasileira de Estudos de População* 1 (1/2): 99-144.
- Massey, Douglas, and Nancy Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67 (2): 281-315.
- . 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Mayer, Susan E., and Christopher C. Jencks. 1989. "Growing Up in Poor Neighborhoods: How Much Does It Matter?" *Science* 243:1441-46.
- Merrick, Thomas W., and Douglas H. Graham. 1979. *Population and Economic Development in Brazil: 1800 to the Present*. Baltimore: Johns Hopkins University Press.
- O'Loughlin, John. 1983. "Spatial Inequalities in Western Cities: A Comparison of North American and German Urban Areas." *Social Indicators Research* 13: 185-212.
- Perlman, Janice. 1976 *The Myth of Marginality*. Berkeley: University of California Press.
- Pfeffermann, Guy Pierre, and Richard Webb. 1979. "The Distribution of Income in Brazil." World Bank Staff Working Paper no. 356. World Bank, Washington, D.C.
- Portes, Alejandro. 1989 "Latin American Urbanization during the Years of the Crisis." *Latin American Research Review* 24:7-44.
- Portes, Alejandro, and Lauren Benton. 1984. "Industrial Development and Labor Absorption: A Reinterpretation." *Population and Development Review* 10:589-611.
- Ribeiro, Luiz Cesar de Queiroz. 1992. "Restructuring in Large Brazilian Cities: The Center/Periphery Model in Question." Paper presented at the meeting of the urban sociology section of the International Sociological Association, Los Angeles, April.
- Sachs, J. 1989. "Social Conflict and Populist Policies in Latin America." NBER

## Socioeconomic Segregation in Brazil

- Working Paper no. 2897. National Bureau of Economic Research, Cambridge, Mass.
- Sassen, Saskia. 1990. *The Global City: New York, London and Tokyo*. Princeton, N.J.: Princeton University Press.
- Schnore, Leo F. 1958. "Social Morphology and Human Ecology." *American Journal of Sociology* 63:620-34.
- . 1965. "On the Spatial Structure of Cities in the Two Americas." Pp. 347-98 in *The Study of Urbanization*, edited by P. M. Hauser and L. F. Schnore. New York: Wiley.
- Shevky, E., and Bell, W. 1955. *Social Area Analysis*. Stanford, Calif.: Stanford University Press.
- Singer, Paul. 1985. *Repartição de Renda: Pobres e Ricos Sob o Regime Militar*. Rio de Janeiro: Zahar.
- Smith, David M. 1979. *Where the Grass Is Greener: Living in an Unequal World*. Baltimore: Johns Hopkins University Press.
- Souza, Paulo Renato. 1980. *Emprego, Salários e Pobreza*. São Paulo: Editora HUCITEC.
- Telles, Edward E. 1992. "Residential Segregation by Skin Color in Brazil." *American Sociological Review* 57:186-97.
- Veja. 1992. "Arruaça na Areia." (October 28).
- Vetter, David. 1988. *Indicadores Sociais*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística.
- White, Michael J. 1986a. *American Neighborhoods and Residential Differentiation*. New York: Russell Sage.
- . 1986b. "Segregation and Diversity Measures in Population Distribution." *Population Index* 52:198-221.
- Yujnovsky, Oscar. 1975. "Urban Spatial Structure in Latin America." Chap. 5 in *Urbanization in Latin America: Approaches and Issues*, edited by Jorge Hardoy. Garden City, N.Y.: Doubleday.



# A Role-Based Ecology of Technological Change<sup>1</sup>

Joel M. Podolny and Toby E. Stuart  
*Stanford University*

This article considers what factors determine whether an innovation becomes a foundation for future technological developments rather than a "dead end." The authors introduce the concept of the technological niche, which includes a focal innovation, the innovations on which the focal innovation builds, the innovations that build upon the focal innovation, and the technological ties among the innovations within the niche. Using patents and patent citations to measure characteristics of innovation niches within the semiconductor industry, the authors show that the size of the niche and the status of the actors within the niche have a positive effect on the likelihood that subsequent innovations will build upon the focal innovation. Competitive intensity within the niche has a negative effect on this likelihood.

## I. INTRODUCTION

The power of technology to mold the arrangement of social activity has long been a topic of sociological interest (Marx 1954; Schumpeter 1950). This concern is manifest in the substantial literature on technology's influence on formal organizational structure (Woodward 1958, 1965; Thompson and Bates 1957; Perrow 1967), informal relations within organizations (Barley 1990), and the survival prospects of individual or groups of organizations (Tushman and Anderson 1986; Barnett and Carroll 1987). The role of technology in shaping population-level outcomes and in structuring organizations has been an intellectually active, central area of inquiry in the sociology-based research on organization-

<sup>1</sup> The authors wish to thank James N. Baron, William P. Barnett, Glenn Carroll, Michael T. Hannan, Jeffrey Pfeffer, and several anonymous AJS reviewers for their comments on this article. Participants in seminars at Cornell University's Johnson Graduate School of Management, the University of Chicago's Graduate School of Business, and the University of California at Berkeley's Haas School of Business Administration provided helpful suggestions on earlier versions of the article. Financial support was provided by the Stanford Graduate School of Business and the Fletcher Jones Foundation. Correspondence may be addressed to Joel Podolny, Graduate School of Business, Stanford University, Stanford, California 94305.

environment relations. However, while great effort has been devoted to the formulation of general propositions regarding technology's ability to configure social relations, there is much less systematic understanding of the dynamic process according to which particular technologies become advanced and extended, while others are never developed. Sociologists have conducted numerous, detailed historical studies that document the direct or diffuse competitions among alternative technologies (Hughes 1983; Law and Callon 1988), but there is a distinct absence of "middle range" sociological theory (see Merton 1968) regarding the determinants of technical change.

Consider any technological domain, such as microelectronics or biotechnology. To the extent that the actors that develop those technological areas devote their energies to one particular pursuit, they will in general be unable to devote those same energies to other undertakings. In effect, ideas and innovations compete with one another for the allocation of resources and attention. Some technological solutions will become "winners" in this competition, drawing the interest and effort of those actors involved in the development of the technological domain. These winners are likely to become a foundation for the future advancement of technological knowledge. Others will be "losers" in this competition in the sense that they will become technological dead ends. While the distinction between winner and loser should be regarded as end points on a continuum rather than as a dichotomy, this distinction helps to frame the central question that this article will consider: What are the factors that determine the degree to which a technology succeeds in this competition for resources and thus becomes a foundation for future technological developments?

If technically superior innovations were always to dominate technically inferior ones, then the answer to this question would depend upon the technical attributes that make a particular innovation superior. However, it is frequently observed that the "best" technologies (e.g., on the basis of price-to-performance ratios) are not necessarily the most successful ones, and this means that technical specifications alone may not be sufficient to gauge the likelihood of technological success (Katz and Shapiro 1984; Farrell and Saloner 1985; Arthur 1988). Because of the uncertainty surrounding the technical promise of an innovation, we argue that an innovation's ultimate contribution to technical advance is not simply contingent on its inherent technical properties, but to a large degree on its "niche."

While we formally define the technological niche of an innovation below, we can anticipate our framework with the observation that most innovations do not emerge in isolation. Rather, they build upon preexisting innovations and may themselves become foundations for future

innovations. Thus, a given innovation is embedded in a relational context that can be defined by its connections to other innovations. Our framework uses the connections among innovations to bound a focal innovation's technological niche. Loosely, we can now state the theoretical claims of this article. We argue that the historical importance of an innovation is largely determined by three attributes of its niche: (1) the extent to which a focal innovation is proximate enough in technical content to the expertise of other innovators that these other innovators are likely to build upon the focal innovation; (2) the competitive intensity, or degree of differentiation, among the innovations to which the focal innovation is linked; and (3) the identities of the actors associated with the focal innovation and the innovations to which the focal innovation is connected. Patents and patent citations, which represent, respectively, innovations and the technological commonalities that indicate technological ancestry, provide the data for an empirical assessment of these claims.

Our article is organized as follows. Section II briefly reviews some of the literature on technical change. Section III develops in greater detail the theoretical underpinnings of the technological niche and also specifies the hypotheses that we test. Section IV describes our data. Section V discusses our study's empirical setting—the semiconductor industry—and our sample—all U.S. semiconductor device patents issued between 1976 and 1991. Section VI presents our measures and our modeling strategy, including the specification of the statistical equations that we estimate. The final two sections present our results and then offer a discussion and conclusion.

## II. TECHNICAL CHANGE IN SOCIOLOGY AND EVOLUTIONARY ECONOMICS

Sociologists and organization theorists have long recognized the importance of technology in affecting organizational and industry-level outcomes, but scholars working in this area have typically been agnostic about the underlying determinants of technical change. To date the sociology-motivated, empirical research on technical change has too often ignored the factors responsible for the genesis of the technologies under study. Thus, in research on technology's influence on social structure, the analyses generally take place after the focal technologies have been developed. Even work that attempts to understand how organizational variables affect whether an invention gains a significant market presence assumes the prior existence of the invention (e.g., Stinchcombe 1990).

One exception to this agnosticism regarding the determinants of technical change is the emerging literature on the sociology of technology, which extends the social constructivist perspective that has been ad-

vanced in the sociology of science (Latour 1987). Through "thick description" of technological controversies and their resolution, the social constructivists emphasize the rich connections among actors and their innovations in technological arenas. These scholars use the network image of a "seamless web" to describe the dense pattern of relations concatenating the innovations and associated actors in unfolding technological systems. Regarding technological progress, a central insight of this view is that the development of new technologies occurs through an inherently uncertain process: There is no archetypal way to accomplish a particular technological endeavor, and therefore there may be many suitable alternatives that satisfy the demands of the technological task at hand (Hughes 1987). This perspective highlights the subjectivity of technological developments; it flatly rejects the notion that technologies are deterministic or that superior innovations necessarily dominate inferior ones (Pinch and Bijker 1987).

Despite the social constructivists' important theoretical contribution, however, a review of this literature raises three concerns. First, it is surprising the degree to which the work in this tradition has proceeded independently of sociological work on organizations and markets (see Tushman and Nelson 1990; Green 1992). This is a concern because the overwhelming majority of innovative activity takes place within the organizational context of what are often industrial firms. Second, the studies in this tradition tend to be retrospective, considering successful technologies and offering *ex post* explanations for their success (Staudenmaier 1985). This methodology causes much of the research in the sociology of technology to suffer from a strong selectivity bias. Third, the use of "network" in this literature has been more metaphorical than rigorous, and, perhaps in part due to the lack of explicit measurement and in part owing to this tradition's preference for thick description, this research has yet to yield generalizable propositions concerning the rate and direction of technical change.

The absence of general, sociological propositions regarding the rate and direction of technical change is noticeable when compared to economic work on this topic. A particularly distinctive aspect of many economic analyses of technical change is the degree to which they might be regarded as sociological in character; many of the assumptions that sociologists find objectionable in mainstream economic theories are significantly relaxed in explorations of technology. For example, the evolutionary approach to the study of technical change (e.g., Nelson and Winter 1982; Dosi 1984) builds upon a conception of the firm developed by behavioral organizational theorists such as Cyert and March. In these models, boundedly rational, "satisficing" (searching for acceptable rather than optimal solutions) organizations follow standard operating procedures in pursuing research and development. This process leads to

a highly path-dependent search for new technologies. The consensus in this literature is that technical change is a gradual process, characterized by the accumulation of minor improvements to incumbent technologies.

As implied by terms such as "technological paradigm" and "technological trajectory," evolutionary theorists contend that technologies develop along paths that are determined by technical properties, as well as problem-solving heuristics and the skills and knowledge contained in a paradigm (Dosi and Orsenigo 1988). Therefore, a simplistic interpretation of evolutionary economic theories of the rate and direction of incremental technical change proposes that these are determined by properties of the technology itself or that they can be inferred from the local search practices of organizations. In these theories, bounded rationality plays a critical role: At the organization level its manifestation is local search, and at the individual level it is apparent in uncertainty about future technological directions. However, such theories lack what we consider to be the defining element of a sociological perspective: a sensitivity to the relational context in which the advancement of technological knowledge takes place.

In the next section, we introduce the concept of a "technological niche" as an analytical construct that focuses directly on the relational context that coevolves with technical change. As the social constructivist perspective does, we highlight the relationships among the actors involved in developing particular technologies. We do so by mapping the technological linkages among innovations that are developed by different actors, and we use the structure and composition of this mapping at one point in time to formulate hypotheses about how it will look in the next period. We rely on patents and patent citations to construct these mappings or to document unfolding technological paths. This approach forgoes some of the thick description that characterizes the work of the social constructivists for a conceptual framework and empirical approach that yields generalizable and falsifiable propositions about the process of technical change.

### III. THE TECHNOLOGICAL NICHE

We define the technological niche at the level of the individual innovation. Using the social constructivist's conception of technical change as an unfolding network as our point of departure, we label a focal innovation's location in this network as that innovation's niche. Each niche consists of nodes, which represent innovations, and ties, which signify the common threads of knowledge that flow from one node to the next as time passes. In our analysis, each innovation occupies an egocentric niche that includes (1) the focal innovation, (2) the innovations on which the focal

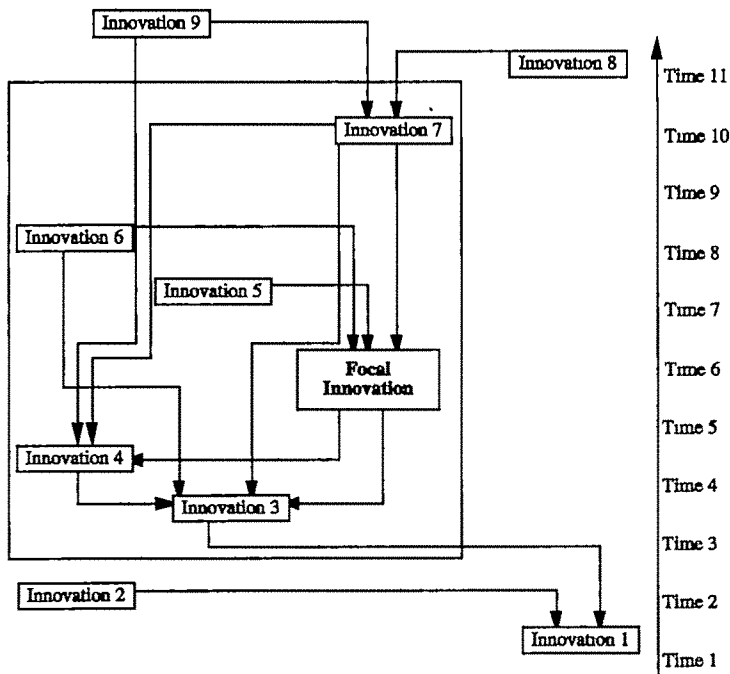


FIG. 1.—Technological niche, defined to include only those innovations to which the focal innovation is directly tied. The box demarcates the boundary of the niche. Arrows indicate patent citations; they point in the direction of the earlier time period, indicating that later innovations build upon earlier ones.

innovation builds, (3) the innovations that build on the focal innovation, (4) the innovations that are sufficiently close to the focal innovation in content that they help to circumscribe the focal innovation's technological contribution, and (5) the technological ties linking these innovations.

Figure 1 provides an example of such a niche. The egocentric niche includes all of the innovations in the box and the technological ties between those innovations. We use the term "tie" to denote technological commonalities among innovations. An innovation A has a tie to B if the contribution of A incorporates, builds on, or is bounded by a technological contribution of B.<sup>2</sup>

<sup>2</sup> Such technological ties between innovations can be regarded as a subset of knowledge-based ties between actors. The broader category of knowledge-based ties would include not only the linkages among innovations but also social ties among actors. Because the actors in our analysis are organizations, examples of knowledge-based ties would include such formal interorganizational alliances as patent cross-license agreements, and technology exchanges. Informal knowledge-based ties would include such connections as personal relationships among technical personnel employed by different organizations.

This egocentric conception of the niche is similar to the conception of the niche as role or relationally defined position (White 1981; Burt 1992; Podolny 1993, 1994; McPherson 1983; DiMaggio 1986). Given this role-based definition, the full technological network can be conceptualized as an interlocking role structure, and each new innovation represents both the emergence of a new niche in that structure and an entrant into one or more established niches. Accordingly, one could analyze technical change either as a process of niche emergence or niche entry. In our view, there is greater analytical leverage to be gained from the latter formulation. We adopt this approach because, if one were to frame the analysis of technical change as a process of niche emergence, it would be unclear where one would look for the causal mechanisms underlying the advancement of technological knowledge. In framing technical change in terms of niche entry, however, the search for the determinants of technical progress turns naturally to an examination of the niche itself, and the central question becomes: How do the characteristics of the niche affect the likelihood of a new entrant?

For our purposes, there are three primary ways to characterize the structure and the composition of the technological niche. One is in terms of the innovations themselves, with purely technical features representing the important distinguishing criteria. The second is in terms of the identity of the actors that own the innovations in a niche. As the social constructivist perspective highlights, underlying every technological network is a community of actors involved in its elaboration. For applied technologies, this community is likely to include business firms with products that incorporate the technologies, as well as universities, government-sponsored research labs, and research consortia that together supply essential technological knowledge to manufacturing firms. The distribution of these actors across the technological network provides a criterion in addition to the technological attributes of innovations to distinguish between niches. The third way to categorize different niches is in terms of niche structure. By structure, we refer to the number and the pattern of relations that connect the innovations in a niche.

Like capital-rich entrepreneurs who might enter one of many organizational niches, innovators may perceive numerous niches into which they might enter.<sup>3</sup> But given fixed resources and perhaps well-established

<sup>3</sup> The question of niche entry can be posed as, Which regions of the technological network should the organization enter? or as, Which product areas should the firm pursue? While there is clearly a strong correspondence between the two, we focus on the technological niche, because nonmarket organizations (i.e., universities and research institutes) play an important role in technological development and because technological development can proceed without influencing markets.

competencies, these actors cannot pursue every possible avenue for future innovation; instead they must adjudicate between niches, deciding to enter only those with the greatest perceived promise. Certain innovations will prove to be central to the advance of technological knowledge and thus of great economic or at least prestige benefit to the organizations that own them. However, others in the same technical arena will remain peripheral to this technological advance. Clearly, the decision to invest resources in a particular niche is one of large consequence to the actor. We now shift our attention to consider how the three attributes of the niche that we have identified are likely to affect the process of technical change.

### The Quality of the Focal Innovation

If the promise of an innovation were an easily observable feature, then the decision of which niche to enter would be a relatively simple one. However, it is striking the degree to which inherent technical properties fail to serve as reliable guides for discerning the innovations that become most successful. For example, when Intel Corporation developed the microprocessor, one of the landmark inventions in the semiconductor industry, the company was unaware of its significance. According to Gilder's (1989) account of the microprocessor's development, at the time of the innovation Intel's sales staff believed that the company would never sell more than 10,000 microprocessors, and the firm's board of directors expressed concern that the chip would distract Intel from focusing on its principal markets.

The observation that technical characteristics alone cannot sufficiently inform decisions about which technologies to develop is a claim that is not unique to academics; it is also made by participants and close observers in different technological domains. For example, a 1993 article in *Electronic Business*, an electronics trade publication, focused on the recent flurry of interest in flash memory, a technology that was important for the inchoate market for personal digital assistants (handheld computers). Flash memory allowed a computer to write and quickly erase information from semiconductors that retained data even after their power supply had been cut. The author writes, "Remember bubble memory? How about Josephson Junction? The chip industry is littered with products and technologies that were the subject of huge amounts of hype but never panned out. Despite these cautionary tales, the drum beats are growing even louder for . . . flash memory. Although flash seems likely to escape the disastrous fates of these earlier technologies, some healthy skepticism seems warranted" (Ristelhueber 1993, p. 99).

We note that pervasive uncertainty is not limited to the development of



specific new product areas. It is often the case that technological debates surround even well-established, basic technologies. For example, a long-standing controversy in the semiconductor industry has pitched silicon against gallium arsenide (GaAs) as the choice material upon which to engrave microscopic circuits (i.e., as the basic material for building integrated circuits). This debate has persisted because silicon has certain properties that make it easier to pack a large number of circuits on a single chip, but GaAs circuits are faster at equal or lower power than silicon circuits. Furthermore, GaAs has properties that make it appealing for building optoelectronic semiconductors (devices that detect, amplify, or transmit light). Still, silicon chips have dominated the history of the industry to date, but for many years some industry participants have foreseen a larger role for GaAs chips. As the former head of IBM's Advanced Gallium Arsenide Technology Laboratory noted, many advocates of silicon referred to GaAs as "the technology of the future, always has been, always will be" (Brodsky 1990, p. 68).

Each of these examples points to a situation in which technical uncertainty surrounded a new innovation, technical controversy, or more generally, a high-technology pursuit. It is precisely because of the uncertainty surrounding predictions of the potential of innovations that the second characteristic of niche composition, the identity of the actors that own the innovation in the niche, becomes important to consider.

#### Attributes of Niche Occupants

We argue that under conditions of technological uncertainty, the status of the actors that sponsor the innovations in a technological niche serves as a tangible guide for the probability that the focal innovation in that niche will become important. We define an innovator's technological status as the perceived quality or importance of that actor's previous contributions to the advancement of technological knowledge. The more that an actor's previous innovations are perceived to have served as the foundation for successful innovational paths, the higher is that actor's status.

In the above quotation on the significance of flash memories, the author expressed skepticism regarding the future promise of this innovation. What is left unexplained, however, is why the "drumbeats are growing even louder" for this technology rather than for the others that the author noted (Josephson Junction and bubble memory). One possible explanation is that the inventor of flash memory technology in 1986 was Toshiba and the leading promoter of this technology was Intel Corporation. Toshiba and Intel are among the most successful and innovative firms in the semiconductor industry. We argue that the combined status of

Toshiba and Intel provided an indication of the likelihood that flash memory would become an important technology.

Anecdotal evidence that a prominent organization or individual actor can bolster expectations for a product by virtue of its association with that product is prevalent. For example, IBM's entry into the personal computer market in 1981 provided the impetus for software developers to devote substantial resources to programs designed for the personal computer (Anders 1981). This example illustrates how the actions of high-status organizations can become "focal points" (see Schelling 1960) for the allocation of resources by the broader array of actors within or around a particular domain.

The general observation that an actor's status affects the attention that the actor receives from the larger community is supported by research in the sociology of science (for a discussion of the connection between the sociologies of technology and science, see Pinch and Bijker [1987]).

Merton (1968) observed how scientists' identities provide a biased lens through which the quality of their intellectual contributions are evaluated. He labeled this phenomenon the "Matthew Effect." Given that the contribution of a scientist's work cannot be assessed until a sufficient amount of time has elapsed for others to develop its implications, a work at its introduction is more readily taken to be important or innovative if it is produced by a high-status member of the scientific community. As a result, the likelihood that a work of a given quality will become the foundation for future research is greater if its author is of higher than lower status.

Sociological research on markets has emphasized a similar status-based dynamic. Specifically, Podolny (1993) argued that a status hierarchy among producers in a market induces a flow of resources that causes a variation in quality across producers' products consistent with, and thus confirming, the initial status ordering. Though Podolny (1993) formulated his argument with respect to the attention and resource allocation processes of consumers, we believe that the argument can easily be extended to the attention and resource allocation processes of innovators.<sup>4</sup> If actors working in a technological area expect that a technology will be superior, they will devote more resources to that technology than to alternatives that are evaluated to be inferior. Consequently, the technologies sponsored by high-status actors are more likely to be rapidly developed than are competing ones, and they will thus appear as superior *ex post* despite the fact that they may not have been superior *ex ante*. For example, if

<sup>4</sup> An interesting empirical question is To what degree is an organization's status in the market related to its status among other innovators? Unfortunately, our data do not allow us to make this assessment.

the combined status of Toshiba and Intel leads others to devote resources to the development of flash memory rather than to competing technologies, such as ferroelectronics, then this additional flow of resources will increase the likelihood that the high expectations for flash memory, as well as the superior status of Toshiba and Intel, will be confirmed.

The fact that an actor's status is expected to lead others to favorably evaluate its innovations does not imply that a high-status actor's innovations will always, or even regularly, be of greater historical significance than those of a low-status actor's innovations. It implies only that on average there will be an *ex post* positive correlation between an actor's status and the acknowledged importance of that actor's innovations. Perhaps of greater importance, the conception of status as an attribution that is contingent on an actor's previous contributions to technological knowledge does not imply that the causal connection between contribution and status is from the former to the latter. To the extent that an actor's status beckons others to enter a particular innovation's niche and thus increases the likelihood that others build upon its innovations, status by definition increases the contribution of an innovation. To draw on another of Merton's expressions, the Matthew Effect may hold in part because status engenders a "self-fulfilling prophecy" with respect to the contribution of an innovation.

In terms of the niche framework, we contend that actors rely on the status of the inventors of focal innovations to inform their decisions about which niches are most likely to become the foundation for superior technologies. These considerations lead to our first hypothesis:

**HYPOTHESIS 1.**—*The greater the status of the actor associated with the focal innovation, the greater the likelihood of a new entrant into the niche at a point in time.*

In addition, we expect that the status of the other actors in the niche will influence the perception of the promise of the focal innovation. To the extent that others use an actor's status to evaluate the importance of the actor's innovation, the actor's status indirectly serves as evidence of the quality of all innovations with which it shares a technological tie. Because we define niches as egocentric networks that comprise clusters of innovations linked by technological ties, we suspect that the status of all of the actors owning the innovations in a niche signals the importance of the central innovation. In effect, when high-status actors publicly acknowledge the importance of an innovation, they draw to it the attention of the other members of the technological community. Accordingly, our second hypothesis states:

**HYPOTHESIS 2.**—*The greater the status of the actors associated with*

*the nonfocal innovations in a niche, the greater the likelihood of a new entrant into that niche at a point in time.*

The discussion of status illustrates certain positive externalities that exist across technological ties. Each innovation (and corresponding actor) helps to legitimate other innovations that possess a similar technological content (and their owners). However, as Schumpeter (1950) highlighted, the process of technical change is in large part a process of competition. Technological ties are a source of legitimation, but also of competition, since they imply a technological commonality among innovations. Accordingly, we suggest that an important factor shaping both the willingness and ability of a would-be innovator to enter a niche is revealed by the structure of that niche.

#### Attributes of Niche Structure

We argue that the relational structure of the niche indicates the intensity of competition within that niche. While competitive intensity is not a directly observable feature of a niche, ecologists (Hawley 1950) have long drawn on Durkheim's (1933) important insight that there is an inverse relationship between differentiation and competitive intensity. More recently, Hannan and Freeman (1989) have emphasized the importance of segregating processes in reducing the competitive intensity in the niche. For our purposes, a crucial analytical question is How can differentiation be measured in a technological niche? Were there some analogue to physical or resource space, then it might be possible to simply count the number of innovations that surround a focal innovation in that space. However, we find implausible the assumption that there is a fixed carrying capacity for a type of knowledge that is independent of the realized level of that knowledge. In the domain of technological knowledge, there is no clear analogue to physical or resource space. Therefore, we rely upon the indirect ties in the egocentric network (i.e., the ties among the innovations to which the focal innovation is connected) to reveal the level of differentiation in the technological niche.

Consider, for example, an innovation that draws on diverse strands of knowledge and then becomes the technological ancestor of a highly differentiated array of innovations. That innovation's egocentric network could be represented as in figure 2. The underlying differentiation is evidenced by the absence of technological connections among the innovations to which the focal innovation is tied. The laser is perhaps an example of such an innovation, because it has spawned multifarious innovations in unrelated domains, such as consumer electronics, medicine, and telecommunications. The fact that the nonfocal innovations do not share

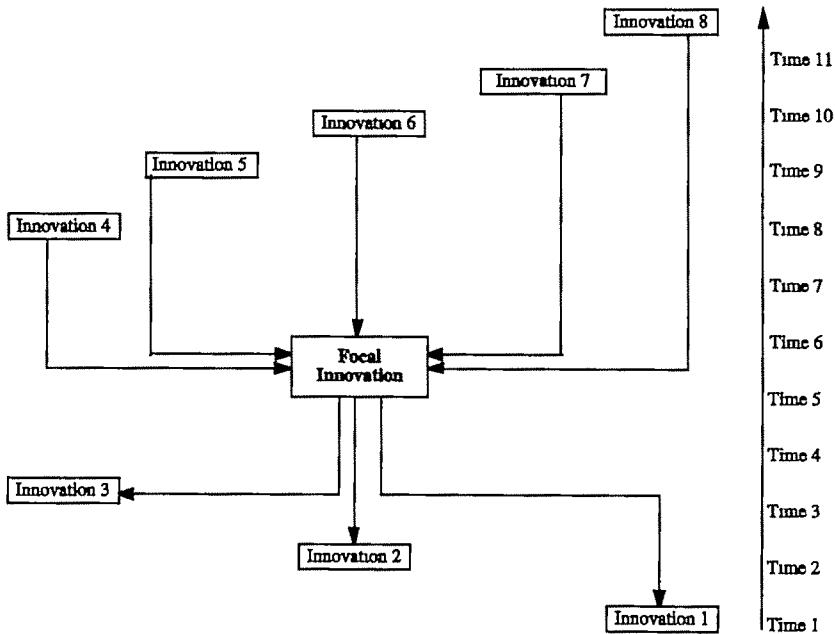


FIG. 2.—Hypothetical sparse niche. The arrows point in the direction of the earlier innovations, indicating that innovations developed later build on earlier ones. The absence of indirect ties among the nonfocal innovations suggests that the figure represents a differentiated niche.

technological ties implies considerable differentiation of technological knowledge in the niche. Conversely, consider the network around an innovation that draws on interconnected innovations and that spawns innovations that are themselves only slightly differentiated. An example of such an innovation, drawn from our own data on the semiconductor industry, is represented in figure 3.

The number of direct ties provides no information about the level of differentiation in the technological content of the innovations within a niche. The central nodes in figures 2 and 3 are both directly connected to an equal number of innovations, and, while the direct ties are relevant to the signaling processes to which we just alluded, it is the indirect ties that provide information about differentiation and hence competitive intensity. The more indirect ties that surround a focal innovation, the more the network can be seen as folding in on itself, blanketing the focal innovation.

The use of indirect ties to measure differentiation can be easily derived from network theory. According to both Granovetter (1974) and Burt

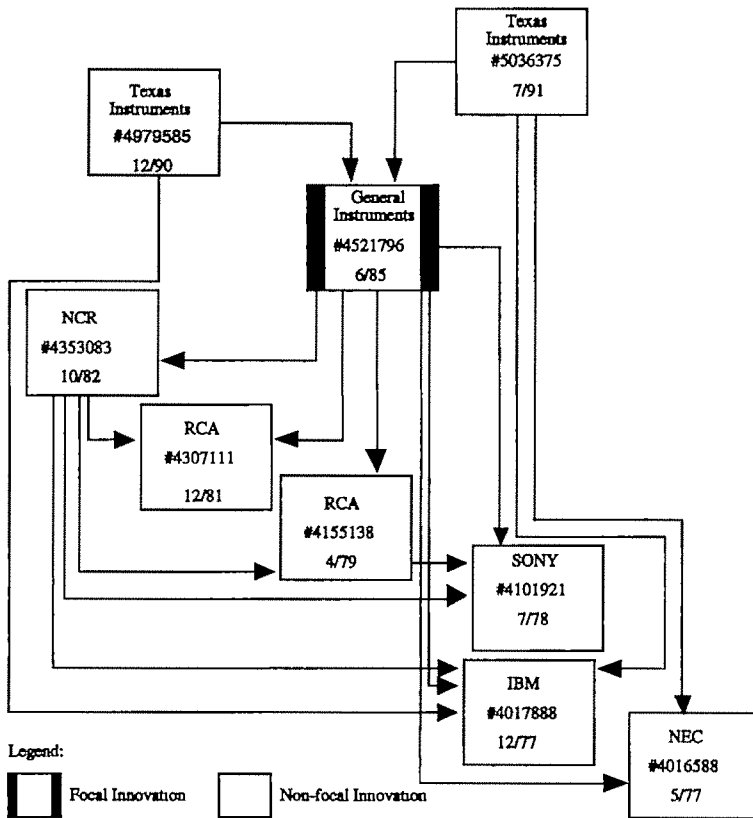


FIG. 3.—Niche of patent no. 4521796. Each box represents a patent and provides information on the patent's owner and issue date. Arrows represent citations from one patent to another. The presence of numerous indirect ties among the nonfocal innovations suggests that there is low differentiation within the niche.

(1992), the information that an actor receives from two alternative contacts is redundant to the degree to which those two actors are connected. More generally, the more connections there are among the nodes within the egocentric network, the less differentiated is the informational content at each node. Thus, the greater the density of technological ties in a technological niche, the more the innovations can be considered to be technologically similar.

These considerations lead to our third hypothesis:

**HYPOTHESIS 3.**—*The greater the number of indirect ties among the innovations in a niche, the lower the likelihood that an innovator will enter that niche at a point in time.*

While we have argued that indirect ties are the crucial structural determinant of the level of competition within the niche, direct ties are still a meaningful attribute of niche structure. Organizational and evolutionary theories of the firm (March 1988; Nelson and Winter 1982) highlight the role of local search behavior in understanding technological evolution. The basic claim of these theories is that resource constraints and embedded organizational routines restrict the areas for innovation that an organization may successfully pursue. This claim can be framed structurally by considering the number of nonfocal innovations, or alternatively the number of direct ties, within the niche. The more that the focal innovation is directly connected to other innovations, the more that the focal innovation is within the local search domain of other members of the technological community. This idea leads to our fourth hypothesis:

**HYPOTHESIS 4.**—*The greater the number of direct ties within the niche, the greater the likelihood of a new entrant into the niche.*<sup>5</sup>

#### IV. MEASURING THE NICHE

Having elaborated the framework and a set of theoretical propositions, we now present a way to operationalize the basic features of the technological niche and the process of niche entry. Within our framework, we require a methodology capable of identifying innovations and the technological connections between them. Our approach is to rely on information contained in patent issues, which grant innovators the legal rights to commercially exploit their inventions. Patents provide a useful means for identifying innovations since they are only granted to products, processes, or designs that are judged by the Patent Office to be industrially useful and nonobvious to those trained in the current state-of-the-art of the relevant technological domain. Patent applications not judged to represent novel technologies are denied.

Patents identify the technological ties between innovations. As part of the application procedure, patentees must list all previously issued U.S. patents that serve as important technological building blocks for the innovations for which they seek approval. Furthermore, it is the role of the

<sup>5</sup> Alternatively, this hypothesis might be formulated in terms of the number of actors within a niche rather than in terms of the number of direct ties. However, we believe that the number of ties is more appropriate, because it takes into account the intensity of a particular actor's search within the area of a focal innovation. We use patents and patent citations to operationalize the niche, and these two quantities (the number of actors and the number of innovations in a niche) rarely differ (see n.8 for the exact proportions of "repeat citations" in the data used for our analysis).

patent examiner to verify that the lists included in the patent application encompass all relevant existing innovations. Such listings are referred to as "prior art" citations, which are an integral part of the patent process in the United States. The citation process is legally important, because it limits the claims of the patent under consideration; the technological domain of the current patent extends only to the point where the prior art ends. The innovator has a legal claim only to the aspects of the patent that do not overlap with the technological contents of the cited patents (see Office of Technology Assessment and Forecast 1976, p. 167).

In our framework, each time a new patent is issued, the patented innovation both enters existing niches (by virtue of overlapping with the technologies represented by previous patents) and represents the emergence of a new egocentric niche. As time progresses, future patented innovations may cite the focal innovation, in which case its niche expands. For example, in figure 1 innovations 3 and 4, which are patented prior to the focal innovation, are included in the niche once the focal innovation is patented. These two innovations represent work that the focal innovation cited as prior art. Innovations 5–7 represent later niche entrants; these innovations cited the focal innovation.

A citation thus designates the focal innovation as a technological precursor to a novel technology. In addition, the Office of Technology Assessment (1976) asserted that the more cites that an innovation receives, the more important that innovation was in the advancement of technological knowledge. The literature has corroborated this assumption. For example, Trajtenberg (1990) found that patent citations were accurate indicators of technological importance in the computed tomography industry, and Albert et al. (1991) found a positive correlation between the number of citations that a patented innovation received and the technological importance that experts ascribed to that innovation. Several cross-sectional studies document a correlation between citations and economic performance. For example, Narin, Noma, and Perry (1987) found high correlations (ranging from .6 to .9) between the possession of a frequently cited patent portfolio and changes in corporate financial measures such as increases in company profits and sales (additional studies are reviewed in Basberg [1987]). In addition to these studies highlighting the relationship between citations and the (perceived) importance of an innovation, the fact that citation analysis is now used by corporations to analyze the technology portfolio of their competitors, to provide insight into likely future market strategies, and to compare productivity within or between firm laboratories further supports the use of patents as proxies for innovative activity (Eerden and Saelens 1991).

While academic research has validated the use of patent citation data



as a meaningful reflection of a number of features of the innovative activities of firms, caution must still be exercised when using patents to identify innovations and the technological ties between them. Innovators may be less willing to patent some types of innovations than others (Levin et al. 1987). Furthermore, there are sometimes industry-specific means that innovators can use to protect their intellectual property.<sup>6</sup> For these reasons, it is not surprising that researchers have found interindustry variance in the tendency to patent (see Scherer 1984). To the extent that innovators do not seek patents for innovations, the use of patents to operationalize the composition and structure of the technological niche will be less encompassing. Clearly, the researcher must be sensitive to the validity of patents as indicators of innovations and the relationships among innovations in the context that is studied. With this admonition in mind, we turn to the empirical setting for our study, the semiconductor industry.

## V. THE SEMICONDUCTOR INDUSTRY

The semiconductor industry began with the invention of the point-contact transistor at Bell Laboratories in 1947. The industry has evolved to include a heterogeneous population of firms, including large captive producers (e.g., IBM),<sup>7</sup> diversified merchant producers (e.g., Motorola), and specialized firms that concentrate on a single technology or market niche (e.g., Bipolar Integrated Technology). European and Japanese organizations have also actively participated in developing semiconductor technologies since the 1950s. In addition, while private firms have conducted a large majority of the innovation in semiconductor technology, national governments and universities have also developed and patented

<sup>6</sup> For example, in microelectronics, the industry that we examine, the Semiconductor Chip Protection Act of 1984 enables innovators to apply for copyrights to protect semiconductor mask works, which are in essence the designs of semiconductor chips' circuitry. However, there are reasons to think that innovators are still likely to seek patents for their devices. First, the period of copyright protection is only 10 years since the time at which the mask work is registered with the Copyright Office or it is first commercially exploited, which is shorter than the 17-year period of patent protection. Further, the Semiconductor Chip Protection Act does not prevent the innovator from receiving both copyright and patent protection for the same invention. Generally, it is thought that the statutory requirements for patents, novelty and nonobviousness, make patents more difficult to obtain than copyrights (for a discussion of the relationship between copyrights and patents for semiconductors, see Ladd, Leibowitz, and Joseph [1986]).

<sup>7</sup> Captive manufacturers are firms that produce for internal use rather than for sale on the open market. IBM is a captive producer because the vast majority of its semiconductors are consumed internally.

technologies. Thus, the semiconductor industry encompasses a range of actors based in many different nations, creating technological networks that span national boundaries and organizational forms.

We have collected all U.S. semiconductor device patents granted to worldwide semiconductor innovators and manufacturers for 16 years, from 1976 to 1991.<sup>8</sup> Previous research suggests that the U.S. patent system is the most complete for analyzing international technology. The United States is widely recognized as the world's largest technological marketplace (with 50% of patents granted to foreign applicants; Albert et al. 1991). Patent information provided by the U.S. Patent Office is available back to 1976 from the Lexis/Nexis on-line database.

As support for the use of patents to examine technology and technological change in the semiconductor industry, we note that all of the landmark innovations in semiconductor technology have been patented (Wilson, Ashton, and Egan 1980). Moreover, during the majority of the time period that we studied, certain semiconductor producers, most notably Intel and Texas Instruments (TI), aggressively litigated to protect their intellectual property. Indeed, the popular press estimated that TI earned some \$1 billion in royalties from infringement lawsuits (Orenstein 1992). TI's patent royalties reached the point that the company reported them as a separate line item on its income statement. In addition, as Japanese producers gained familiarity with the U.S. legal system, they too became more assertive in filing infringement cases (for example, Fujitsu recently opted to file a countersuit against TI rather than to pay royalty fees; Helm 1992). The apparent importance that semiconductor firms placed on patents and patent rights suggests the validity of using patent information as operationalizations of the central concepts of our analysis.

## VI. ANALYSIS

To assess our hypotheses, we model the citation rate for existing patents or, in theoretical terms, the process of niche entry and hence extension. Each spell begins when a patent is issued or cited and ends when it is next cited.<sup>9</sup> We employ the proportional hazards model introduced by

<sup>8</sup> Under the U.S. system, patents are filed according to major class and subclass. Semiconductor device patents include all subclasses of primary class 357. Patents are typically filed in one primary class/subclass combination but are also cross-classified in additional locations. We include in our sample all patents filed in primary class 357.

<sup>9</sup> We experimented with alternative event definitions. For example, we performed analyses in which we did not end a spell when a self-citation occurred (i.e., when a focal patent was cited by its owner). This event definition can be defended on the grounds that a self-citation is substantively less meaningful than a citation from another actor, because it does not imply that an additional actor has decided to enter

Cox (Kalbfleisch and Prentice 1980). The equation that we estimate takes the following specification:

$$r(t) = h(t) \exp[XB + Y(t)S], \quad (1)$$

where  $r(t)$  is the transition rate or hazard of niche entry,  $h(t)$  is an unspecified baseline rate for the transition,  $X$  is a matrix of time-constant covariates,  $Y(t)$  is a matrix of time-varying covariates, and  $B$  and  $S$  are vectors of unknown regression parameters. Because  $h(t)$  is an unspecified step-function, the Cox model offers an extremely flexible means for modeling time dependence. While the Cox model accounts for interarrival time dependence (i.e., the time from when the patent is granted to the time at first entry and then the time between subsequent entries) with an unspecified baseline rate for the transition, the time since last arrival is not the only form of time dependence that is likely to affect the rate of citation. In addition to modeling interarrival times, we also include two clocks that accelerate the baseline rate. First, we include a variable (updated monthly) to denote the calendar time. Second, assuming that firms are more likely to enter a niche after they become aware of it, the composite baseline should increase as a function of the time since the patent at the center of the niche was introduced. We therefore include a variable in the model to denote the time that has passed since the introduction of the focal patent (the age of the patent, also updated monthly). However, because we expect the relevance of a patent to decrease with the time since it was introduced, we include an age-squared variable as well to allow for nonmonotonicity.

Our theoretical framework emphasized three attributes of the technological niche: the quality of the focal innovation in a niche, the attributes of the actors associated with the niche, and the structure of relations among the innovations in the niche. We now introduce the covariates that we include in the model to represent each of these features of the niche.

*The quality of the focal innovation.*—The difficulty that industry observers have in inferring the comparative quality of innovations almost necessarily implies the absence of a completely adequate means to control

---

the niche. Based on a similar logic, we considered a second alternative in which a spell ends only if the focal patent is cited by a new actor rather than one that has previously cited the focal patent. Again, it might be that a citation from an additional actor is a more significant event than a repeat citation, because it indicates that an additional actor has chosen to enter the niche. However, only 5.3% of the spells end in self-citations, and only 4.6% of the spells end in repeat citations by nonfocal actors. Consequently, while there may be theoretical reasons for considering alternative event definitions, in practice the results were unaffected by the different definitions of the dependent variable.

for the quality of an innovation. We see three possible ways to respond to this difficulty. One is to forgo any attempt to control for quality, under the assumption that what is not observable cannot have an effect on the likelihood of niche entry. A second approach is to treat quality differences across patents as unobserved heterogeneity and then to devise some method for controlling for this unobserved heterogeneity. A third possibility is to rely on one of the measures of quality that has been put forth in economic analyses of patents as indicators of innovative activity. In our analysis, we have chosen the second and third approaches, which both suggest the same control variable for quality.

In their discussion of unobserved heterogeneity, Heckman and Borjas (1980) noted that unobserved differences across units are likely to result in occurrence dependence. A frequently cited example of this type of unobserved heterogeneity arises in the context of research on job mobility. If there is unobserved heterogeneity across individuals in their likelihood of shifting jobs, then an occurrence-dependent term (i.e., the number of times that someone has changed jobs in the past) will have a positive effect on the rate of job mobility. Heckman and Borjas argued that one way to account for such unobserved differences is to include as a control variable the number of previous realizations of the dependent variable. Applying Heckman and Borjas's logic to our analysis of patents, we include as a covariate the number of times that a patent has been cited, controlling for the time that the patent has been at risk of being cited.

While econometric work on unobserved heterogeneity suggests the inclusion of a variable denoting the number of times that the patent has been cited, so too does economic research on patents as indicators of innovative activity. In particular, Trajtenberg (1990) employed the number of citations received by a patent over a given time interval as a measure of that patent's quality. By including the number of citations received as a covariate (while also controlling for time dependence in the form of calendar time, the age of the patent, and the age of the patent squared), we thus include as a covariate the measure of quality that has been put forth in the economics literature. While we wish to avoid equating an occurrence-dependent term solely with quality differences, since unobserved heterogeneity could be due to other factors, the fact that we have included the measure that others have used to capture quality should serve as a counter to the alternative hypothesis that many of our results can be explained by unobserved quality differences between patents.

*Attributes of niche occupants.*—We have argued that an important characteristic of the actors within a niche is their status. Recall that we define an actor's status at a particular point in time as its contribution

to the advancement of technological knowledge up to that time. Consistent with this definition, we measure an actor's status as a proportion related to the number of citations that the actor has received on all of its semiconductor device patents during the 12-month interval prior to the month in which it introduces a patent. A greater score on this variable implies a greater contribution to the advancement of technological knowledge over that time period. We do not measure status as a simple count, because an actor's portfolio of patents enlarges over time, thus increasing the number of patents "at risk" of a cite. This procedure would result in status increasing merely as a function of time. For example, in month 16 an actor can receive citations on patents only during a 15-month time period, but in month 36 it can receive citations on patents that it issued over a 35-month time period. To correct for the expanding risk set, we divide the number of cites to an actor over a given 12-month interval by the cites to all actors during that window. As a consequence of this standardization, an actor's status can range from "0," if it has received no citations during a given window, to a maximum of "1," if it has received all of the citations made over the previous year. In the context of our data, the highest status organizations have scores of approximately .1, indicating that they had received 10% of the citations made over a one-year period.

Hypothesis 1 anticipated a positive relationship between the status of the focal organization (i.e., the owner of the focal patent) and the rate at which other organizations choose to enter its niche. We thus include as a covariate the status of the focal organization at the time of its most recent entry into the focal patent's niche. In general, the value of this covariate will be the focal organization's status at the time that the focal patent was approved. However, if the focal inventor reenters the niche by introducing a patent that cites the focal patent, we update the variable to reflect the focal actor's status at the time of its most recent entry into the niche.<sup>10</sup>

To clarify the definitions of the variables in the event-history analysis, we refer the reader to figure 4 and table 1. Figure 4 presents a hypothetical niche that expands over time. Table 1 presents the spell data that would correspond to such a niche. The first spell begins when the focal patent is granted to organization C. At this time, organization C has a status of .04, and thus the covariate denoting the status of the focal inventor at the time of its most recent entry into the niche takes the value

<sup>10</sup> Clearly, this is not the only coding rule that one might apply. One might, for example, update the focal actor's status whenever it changed. However, such a coding rule was impractical, because many actors change status monthly as new citations arrive and old citations are dropped (these are typically small changes).

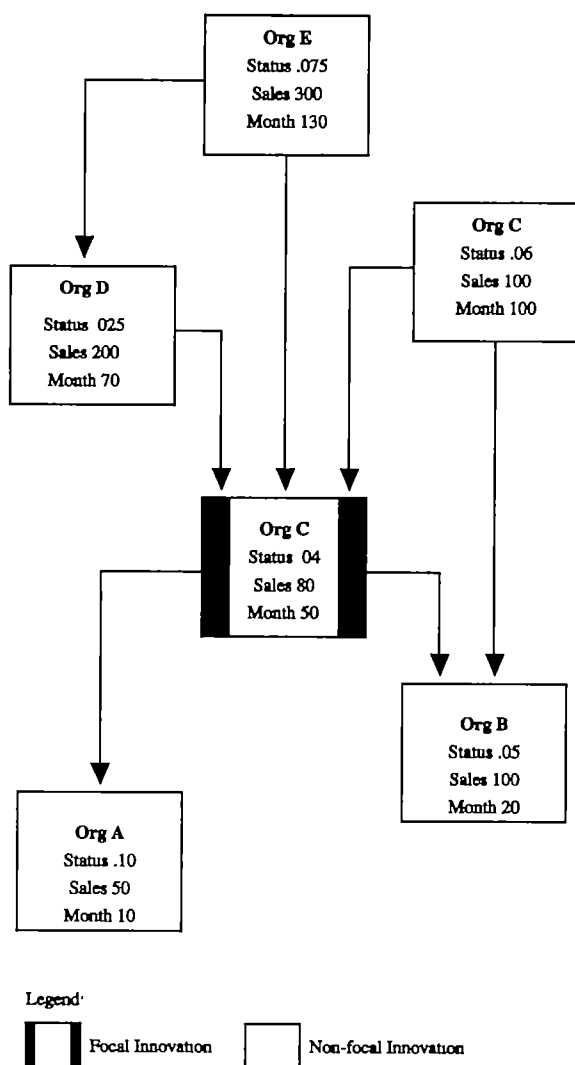


FIG. 4.—Hypothetical niche and attributes of organizational actors from which the spells in table 1 are constructed. Each box represents a patent and provides information on the month in which the patent was issued as well as the status and sales of the patent's owner at that month.

TABLE 1

SPELLS CONSTRUCTED FROM THE HYPOTHETICAL NICHE IN FIGURE 4

VARIABLES	SPELLS			
	1	2	3	4
Time clocks *				
Age of patent . . . . .	0	20	50	80
Age of patent squared . . . . .	0	400	2,500	6,400
Calendar time . . . . .	50	70	100	130
Unobserved quality of the focal patent:				
<i>N</i> of patents that cite the focal patent . . . . .	0	1	2	3
Attributes of actors within the niche:				
Status of focal organization . . . . .	.04	.04	.06	.06
Average status of nonfocal organizations whose patents are cited by the focal patent . . . . .	.075	.075	.075	.075
Average status of nonfocal organizations that cite the focal patent . . . . .	0	.025	.025	.05
Sales of focal organization . . . . .	80	80	100	100
Average sales of nonfocal organizations whose patents are cited by the focal patent . . . . .	75	75	75	75
Average sales of nonfocal organizations that cite the focal patent . . . . .	0	200	200	250
<i>N</i> of alliances entered into by the focal organi- zation . . . . .	0	0	0	0
<i>N</i> of self-citations by the focal organization . . . . .	0	0	1	1
Attributes of the niche structure				
<i>N</i> of indirect ties . . . . .	0	0	1	2
<i>N</i> of patents cited by the focal patent . . . . .	2	2	2	2

\* Values listed are at the start of the spell

.04. This value implies that in the 12 months prior to month 50, the month in which the focal patent is introduced, organization C received 4% of all the citations made. The first spell ends when the focal patent is cited by organization D in month 70. The covariate for the status of the focal inventor remains the same for the duration of the second spell, which ends in month 100 when organization C reenters the niche with a self-citation. On account of this reentry, we update the value of the covariate of the focal inventor to reflect the fact that its status has changed from .04 to .06. This variable has a value of .06 for the duration of the third and fourth spells.

Hypothesis 2 posited that the status of the nonfocal inventors in a niche would have a positive effect on the likelihood of new entrants into that niche. To assess the effect of the status of the nonfocal actors, we include two covariates: the average status of the nonfocal innovators with patents that are *cited* by the focal patent and the average status of the nonfocal innovators with patents that *cite* the focal patent. We divide the status of the nonfocal innovators into these two components because

of the asymmetry implicit in the patent citation process. Our theoretical framework suggests that a citation from a high-status actor ought to be a more meaningful event than a citation to a high-status actor. If a high-status actor cites a focal patent, that actor implicitly acknowledges the dependence of its innovation on the technologies represented by the focal patent. In contrast, if a focal patent cites the innovation of a high-status actor, this implies only that the focal patent is in a technological area that the high-status actor regards as important; such a cite from the focal patent to the innovation of a high-status actor does not imply any acknowledgment on the part of the high-status actor of the focal patent's contribution.

Once again, figure 4 and table 1 show how these two variables are constructed. The focal patent in figure 4 cites two innovations. The average status of the owners of these patents is  $(.10 + .05)/2 = .075$ . The average status of the nonfocal innovators cited by the focal patent maintains this value for all of the spells. In contrast, the average status of the nonfocal innovators that cite the focal patent changes each time an additional organization enters the niche. In the first spell, the focal innovation by definition has no citers, and thus the value of this variable is 0 until the focal innovation is cited by organization D. Since organization D's status is .025, this variable has a value of .025 for the second spell. The value remains the same for the third spell since the reentry of the focal innovator into the niche does not affect this variable. However, when the focal patent is cited by organization E, this variable changes in value to .05, the average of organization D's status and organization E's status.

While status is the actor attribute of greatest theoretical interest, we recognize that it is not the only one that may be relevant to the process of niche entry. Another important attribute is the market presence of the organizations in the niche. Just as an organization's status may be a signal of the technological significance of its innovations, so sales may be a signal of the market significance of its innovations. There may be other reasons why an organization's market presence would affect the rate at which its technologies are developed, but we are less interested in discriminating among alternative explanations than we are in including market presence as a control. Because it seems reasonable to believe that sales and status will be correlated, greater confidence can be placed in a status effect if it can be shown to have explanatory power in excess of an organization's market presence.

Motivated by these considerations, we admit in the analysis the market sales volumes (in billions of dollars) for the merchant and captive semiconductor makers for which we were able to obtain this information. Data on the annual sales of merchant semiconductor firms between 1981



and 1991 were supplied by Dataquest, a consultancy and market research firm with clients in high-technology industries. Data on the estimated dollar value of production volume for the major captive producers over this period were obtained from the annual status reports issued by the Integrated Circuits Engineering Corporation (McLean 1981-91). Like our measures for organizational status, our measures for organizational sales are based on the time of an organization's most recent entry into the niche. However, unlike our status measures that are updated monthly, the sales data are available only annually. Thus, we include as a variable the sales volume of the innovator of the focal patent in the year prior to its most recent entry into the niche. We also add variables representing the average sales of all of the nonfocal organizations that are cited by the focal patent and the average sales of all the nonfocal organizations that cite the focal patent. Each time a new innovator enters a niche, this variable is changed to reflect the sales of the entering organization (measured at the year prior to entry).

There are two complications with the sales data. First, while the Dataquest data account for over 90% of worldwide semiconductor sales and offer the most complete information available, they are not exhaustive. Dataquest attempts to collect data on all firms producing more than \$10 million in sales (approximately .02% of the market in 1991), but it does not track sales for all firms producing less than this dollar amount. For firms not included in the database, we assign a sales volume of zero. In so doing, we assume that sales levels of less than \$10 million are not a stronger signal of market potential than zero sales, such as would be true of nonproducing assignees (e.g., universities, government agencies, etc.). As justification for this decision, we note that the total sales volume in the industry ranged from \$15.9 billion in 1981 to \$55.5 billion in 1991, and the mean sales for a firm in the database are \$473 million, so zero sales seems a sufficiently close approximation for the excluded firms. The second complication is that, while we have semiconductor patents from 1976, we have data on sales volume only from 1981. Excluding spells affected by missing data results in the elimination of 47% of the sample, quite possibly biasing the results. Therefore, we imputed missing sales data according to the multiple-imputation procedure advocated by Rubin (1987).<sup>11</sup>

<sup>11</sup> A distinctive feature of Rubin's multiple imputation procedure is that it introduces a random component into the imputation to reflect the uncertainty about the "true" parameter values underlying the coefficients that one uses for the imputation. In our analysis, we used a regression model to impute the missing sales data in year  $t - 1$  from observed (or imputed) sales in year  $t$ . We repeated this procedure until we had a complete data set (i.e., complete sales information for all firms going back to 1975).

We include two additional control variables that reflect other aspects of the actors in a niche. First, we include a variable that is the number of times that the focal organization cites itself. We noted earlier the possibility that a citation by the innovator of the focal patent might not represent as significant an event as a citation by an independent firm. Thus, we control for self-citations in the analysis. The final attribute of the actors in the niche that we include is the total number of patent license, patent cross-license, technology exchange, and second-source agreements that the focal organization entered prior to its most recent entry into the niche.<sup>12</sup> We include the number of strategic alliances as a control variable, because the types of interorganizational relationships captured in this variable are likely to affect the degree to which the technology owned by an organization diffuses.<sup>13</sup> Specifically, we expect that an organization's participation in interorganizational relationships increases the amount of overlap between its innovative efforts and those of the organizations with which it is partnering. This overlapping effort may increase the degree to which other organizations build on the focal organization's patents. The data for this variable were collected by scanning every edition of the trade publication *Electronic News*, a weekly periodical that has a section devoted to the semiconductor industry, for the time period covered by the focal patents in our study. *Electronic News* is probably the most comprehensive single source for information about collaborations among semiconductor firms.<sup>14</sup>

*Attributes of niche structure.*—While hypotheses 1 and 2 considered the effects of status on niche entry, hypotheses 3 and 4 were concerned with the relationship between the structure of the niche and the rate

---

We then constructed another complete data set following the same procedure. Rubin (1987) explains how parameter estimates from two or more such complete data sets can be combined to provide unbiased estimates of coefficients and their variances.

<sup>12</sup> Second-source arrangements are a common practice in the semiconductor industry. A formal second-source agreement occurs when one firm licenses another to manufacture one of its products, resulting in a nearly identical copy of the licensor's design. This practice helps to guarantee users that the product will be reliably supplied.

<sup>13</sup> Some of these agreements (cross-license and technology exchange deals) are symmetric in the sense that each firm licenses their patents or exchanges their technology with the other. The others, license and second-source deals, are asymmetric, because one firm is the licensor and the other is the licensee. Because we are interested in controlling for other factors that may lead to the diffusion of an organization's technology (and hence an increase in the rate at which its patents are cited), we increment this variable for both firms in the case of symmetric deals but only for the licensor in the case of second-source and license deals.

<sup>14</sup> While *Electronic News* is the most comprehensive source for interfirm collaborations, it is unlikely that the data on alliances is complete. In particular, we suspect that we have missed partnerships between small firms and especially small firms based outside the United States.

of entry. Specifically, hypothesis 3 stated that indirect ties among the innovations in a niche are expected to have a negative effect on the likelihood of a new entrant. The argument is that the competitive intensity in a niche is a positive function of the number of indirect ties in that niche, controlling for the number of direct ties. Therefore, we include as a covariate in the analysis a count of the number of indirect ties among the innovations in a niche (see fig. 4).

Hypothesis 4 speculated that the number of direct ties in the niche would have a positive effect on the likelihood of a new entrant into the niche. As with the sales and the status of the nonfocal innovators in the niche, direct ties can be decomposed into two components: (1) the number of patents that the focal innovation cites and (2) the number of patents that cite the focal innovation. The previous discussion of quality differences across patents suggests that there is an identification problem encountered in interpreting the number of citations received by the focal innovation. While the number of citations to the focal innovation could represent the degree to which the focal innovation is in the local search domain of other actors, it could also be a proxy for unobserved heterogeneity or patent quality differences. Thus, this parameter cannot be given a unique interpretation. While the interpretation of an effect for the number of citations *going to* the focal patent is confounded by alternative processes that may be proxied by this quantity, the number of citations *made by* the focal innovation can be clearly interpreted. The more patents that the focal innovation cites, the more that the focal innovation is within the local search domain of those within the technological community. Accordingly, we include the number of citations made by the focal patent as a test of the hypothesis 4.

We reiterate that our sample includes all semiconductor device patents between 1976 and 1991. While these data provide the opportunity for a longitudinal analysis, it is important to point out that the arbitrary lower bound of 1976 leads to left censoring. When a niche emerges, it includes not only the focal patent but also all of the patents that the focal patent cites. If any of these patents were issued prior to 1976, they would be excluded from the data and thus not captured in the demarcation of the focal patent's niche. To minimize the potential biases from left censoring, semiconductor device patents issued between 1976 and 1981 were not included in the analysis as focal patents but only as potential cites of patents issued from 1982 to 1991. Given that the product life cycle in the semiconductor industry is typically estimated to be three to five years from introduction to maturity (McClean 1981–91), we believe that a six-year window is fairly conservative. We have experimented with the length of this window, and the results that we report are robust across shorter and longer intervals. In total, our sample contains 4,048 patents,

TABLE 2

## FREQUENCY OF EVENTS PER PATENT

<i>N</i> of Events	Frequency
0	1,037
1	668
2	483
3	289
4	174
5	111
6	72
7	32
8	34
9	25
10+	58

NOTE.—An event occurs when a subsequent patent enters a focal patent's niche, each event constitutes a completed spell

TABLE 3

## DESCRIPTIVE STATISTICS FOR VARIABLES

Variables	Mean	SD	25th Percentile	75th Percentile
Quality of the focal innovation:				
<i>N</i> of cites received by focal patent	1.743	3.092	0	2
Attributes of niche occupants				
Status of focal innovator	.030	.024	.005	.048
Average status of nonfocal innovators cited by the focal patent	.020	.020	0	.033
Average status of nonfocal innovators that cite the focal patent	.013	.019	0	.026
Sales of focal innovator	913	1,246	0	1,361
Average sales of nonfocal innovators that are cited by the focal patent	335	.553	0	415
Average sales of nonfocal innovators that cite the focal patent	.723	1.132	0	1.233
Self-citations of focal innovator	.453	.923	0	1
Alliances formed by focal innovator	5.789	6.976	1	8
Attributes of niche structure.				
<i>N</i> of citations made by the focal patent	2.278	2.142	1	3
<i>N</i> of indirect ties				
Niche size = 2 or 3 ( <i>N</i> = 3,244)	.164	.440	0	0
Niche size = 4 or 5 ( <i>N</i> = 2,154)	.617	1.010	0	1
Niche size = 6 or 7 ( <i>N</i> = 1,171)	1.432	1.741	0	2
Niche size = 7 or 8 ( <i>N</i> = 645)	2.403	2.592	0	3
Niche size ≥ 9 ( <i>N</i> = 753)	5.678	5.269	2	8

of which 2,983 represent focal patents (i.e., were issued after 1982). Table 2 presents descriptive information on the number of entries (completed spells) per patent. For example, 483 patents experienced two complete spells (i.e., two subsequent patents entered their niche). Table 3 presents descriptive information on critical independent variables. In addition to the means and standard deviations, we present values for the 25th and 75th percentiles, because the interquartile range is less sensitive than the standard deviation to outliers and to the assumption of normality. We report the distribution for the number of indirect ties at various levels of niche size (i.e., the number of citations from the focal patent plus the number of citations to the focal patent), since the number of possible indirect ties is contingent on the number of direct ties.

## VII. RESULTS

We estimate the waiting time model specified in equation (1) using TDA 5.2 (Rohwer 1993). Hazard rate estimates are presented in table 4. The coefficients reported in this table indicate how a one-unit change in an independent variable serves to multiply the rate of niche entry.

In columns 1 and 2 of table 4, the effect for the status of the owner of the focal patent is positive and statistically significant. However, when the attributes of the nonfocal innovators and the sales of the owner of the focal patent are included in the full model (col. 3), this effect is no longer statistically significant. We had suspected that the lack of significance may have resulted from a high correlation between the sales and status of the owner of the focal patent, but the correlation between these two variables is only .55. Thus, we cannot reject the null hypothesis that the status of the focal innovator has no independent effect on the likelihood of a new entrant into the niche of the focal patent.

In contrast, the average status of the nonfocal innovators that cite the focal patent and the average status of the nonfocal innovators that are cited by the focal patent both have positive and statistically significant direct effects on niche entry in the complete model. These findings indicate that the relationship between organizations in a niche can be commensalistic: When patents owned by high-status actors enter a niche, they attract other innovators, thus enhancing the status of the focal actor. Moreover, the status of those that cite the focal patent has a greater effect on the hazard of niche entry than does the status of those who are cited by the focal patent. To illustrate the difference, we consider how an interquartile shift in both variables multiplies the rate of entry. An interquartile-size increase in the average status of the citers of the focal patent augments the rate of niche entry by 30% ( $\exp[10.14 \times .026] = 1.302$ ). In contrast, an interquartile-size increase in the average status of

TABLE 4  
PARAMETER ESTIMATES FOR THE HAZARD OF NICHE ENTRY

Variable	Model 1	Model 2	Model 3
Time clocks:			
Age of patent	.0549** (.0058)	.0458** (.0058)	.0453** (.0058)
Age of patent squared	-.0001** (1.6e-5)	-.0001** (1.8e-5)	-.0001** (1.8e-5)
Calendar time	-.0125** (.0006)	-.0086** (.0007)	-.0089** (.0007)
Quality of the focal innovation:			
N of cites received by focal patent	.1118** (.0064)	.1030** (.0065)	.1019** (.0066)
Attributes of niche occupants:			
Status of focal innovator	1.3937* (.6555)	1.6508** (.5934)	.7780 (.6605)
Average status of nonfocal innovators cited by the focal patent	...	2.1555** (.8239)	2.5329** (.8326)
Average status of nonfocal innovators that cite the focal patent	...	10.0709** (1.4007)	10.1462** (1.4033)
Sales of focal innovator	0.0292* (.0148)	...	0.0474** (0.0151)
Average sales of nonfocal innovators cited by the focal patent	...	.1710** (.0332)	.1644** (.0332)
Average sales of nonfocal innovators that cite the focal patent	...	.0760** (.0230)	.0757** (.0231)
Self-citations of focal innovator	-.0006 (.0165)	.0382* (.0163)	.0276 (.0168)
Alliances formed by focal innovator	.0051* (.0023)	.0037 (.0022)	.0022 (.0023)
Attributes of niche structure:			
N of indirect ties	-.0557** (.0078)	-.0467** (.0080)	-.0464** (.0080)
N of cites made by focal patent	.0933** (.0066)	.0692** (.0071)	.0734** (.0072)

NOTE.—SEs in parentheses

\*  $P < .05$

\*\*  $P < .01$ .

those cited by the focal patent increases the rate of niche entry by only 8.7% ( $\exp[2.53 \times .033] = 1.087$ ). In other words, the effect of the average status of those who cite the focal innovation is more than three times the magnitude of the effect of the average status of those cited by the focal patent. The difference in the magnitudes of these coefficients is consistent with our interpretation that a citation to the focal innovation is an explicit recognition of the importance of that innovation and that a citation from the focal innovation indicates only that the focal innovation is proximate to the innovations of other actors.

Particularly interesting is the finding that the status of the nonfocal actors has a greater direct effect on the likelihood of niche entry than does the status of the focal innovator. Even in the model for which the sales of the focal organization are excluded from the analysis (col. 2 of table 4), the effect for the status of the nonfocal actors is greater than the effect for the status of the focal actor. While we had not anticipated this result, it is compelling to find that innovators have greater difficulty legitimizing their own innovations than drawing attention to the innovations of others. An actor's self-interest in the success of its own innovations almost undoubtedly compromises its ability to draw on its status for the purpose of attracting to its innovations the attention of other actors.

Although there is no statistically significant direct effect for the status of the focal innovator on the likelihood of entry, a comparison of the results in columns 1 and 3 of table 4 reveals an important indirect effect of the focal organization's status. In column 1, when only the attributes of the nonfocal innovators are excluded, the effect of the focal owner's status on the likelihood of niche entry is positive and statistically significant. Indeed, the magnitude of the coefficient is approximately two times greater than in the full model. This finding implies that the status of the focal innovator significantly affects the probability that higher status (and larger) nonfocal actors will enter the niche, which in turn has a positive effect on the rate of niche entry. However, when we control for the status and size of these nonfocal actors, there is little direct effect of the focal owner's status. Together, the indirect effect for the status of the focal actor and the direct effects for the nonfocal actors underscore the importance of the sociotechnical context that is highlighted by the niche framework. The less that an actor is able to exploit its status to draw attention to its own efforts, the more that actor depends on the context in which its innovations are situated.

Now we turn to the results for the attributes of niche structure. As hypothesized, the parameter estimate for the number of indirect ties in a niche is negative, indicating that an increase in the number of indirect

ties reduces the rate of niche entry. Controlling for the number of direct ties in the niche, each indirect tie lowers the rate of niche entry by 4.5% ( $\exp[-0.0464] = .955$ ). The negative coefficient as well as the level of significance for this variable suggest that the number of indirect ties does reflect the competitive intensity in the niche. Thus, technological domains that are "crowded" in the sense that the innovations in them lack differentiation are ones that potential entrants either avoid or are unable to enter.

Finally, the coefficient for the number of cites made by the focal patent has a positive effect on niche entry. This finding supports the local search hypothesis, which suggests that organizations continue to work in the domains in which they have had previous successes. The more that an innovation is within the local search domain of other organizational actors, the more likely is that innovation to become a foundation for future technological developments.

Turning briefly to the control variables, we find that the coefficients for the number of cites received by the focal patent is positive. Because this variable is an occurrence-dependent term, we urge caution in its interpretation. However, the fact that this measure has been used in the economics literature to control alternatively for unobserved heterogeneity or quality differences across patents should increase the level of confidence in the results. The estimate for the sales volume of the focal firm and the average of the sales volumes of the nonfocal firms in a niche both have positive effects on the rate of niche entry. The number of times that a firm cites its own patents does not have a statistically significant effect in the complete model. The final control variable, the number of strategic alliances in which a firm participates, is positive but not statistically significant in the complete model.

## VIII. CONCLUSION

The sociology of technology has provided insight into the process of technical change by emphasizing the uncertainty inherent in the innovative process and by invoking the network metaphor to highlight the interconnections among the diverse array of actors involved in the development of technologies. However, as we have observed, a review of this literature raises three concerns. First, it is unfortunate the degree to which this literature has developed separately from the sociological work on organizations and markets, given that the vast majority of innovation takes place within (typically market-based) organizations. Second, the studies in this tradition tend to be retrospective and thus suffer from a selectivity bias. Third, while this tradition's preference for "thick description" has



sensitized scholars to many intricate features of the innovative process, this description has often proceeded at the expense of generalizable, falsifiable propositions regarding the rate and direction of technical change.

In this work, we have sought to redress these shortcomings. While we have depended upon the social constructivists' emphasis on the inherent uncertainty underlying the innovative process and the "seamless web" connecting actors and innovations, our framework and hypotheses have drawn on insights from the organizations literature within sociology. Ecology and especially network theory have provided the tools to operationalize the technological niche, and, along with recent sociological work on markets, they have contributed to the formulation of a series of general hypotheses about how the properties of a niche and the actors within that niche affect the likelihood of entry.

In the hazard rate analysis, we found that the status of the focal actor did not have a statistically significant direct effect on the rate of niche entry, but it did exert an indirect effect. High-status focal actors were more likely to find their innovations in niches occupied by high-status nonfocal actors, and the status of these nonfocal actors increased the rate of niche entry. Since an organization's status derives from the degree to which others have entered the niches in which it is the focal innovator, this result indicates that an actor's status depends not only on the quality or importance of its past efforts but on the status of those with whom it is affiliated. This finding and the finding that the number of indirect ties in the niche (which proxied for competitive intensity) had a negative effect on the rate at which innovators enter are the ones that we regard as the most interesting. These results demonstrate that an organization's role in the process of technological change is meaningfully embedded in the sociotechnical context into which its innovations are introduced and developed.

Before concluding, we would like to draw attention to some features of this analysis that have so far remained only implicit, as well as to highlight some directions for future research. Embedded in our analysis has been a methodological proposition about the process of technical change. In the literature on technology, the distinction is often made between the rate, direction, diffusion, and adoption of innovations, and these phenomena have been explored in separate studies. While we acknowledge the conceptual separation of these areas, we do not subscribe to their analytical divorce. Technologies develop as they diffuse, and as they progress they become more attractive to potential adopters, affecting the pace at which the initial innovation is modified; thus, rate, direction, diffusion, and adoption are intertwined.

While this analysis has been carried out at the microlevel, we briefly comment on the macrolevel implications of the observed effects of status

and sales insofar as they indicate possibilities for future research. If large and high-status organizations shape the course followed by technical change, then within any technological domain we should observe that as history progresses innovative activity will be increasingly conducted by large, high-status firms, since small, low-status firms will be more likely to find themselves following technological dead ends. Indeed, Schumpeter (1950) posited a similar macrolevel dynamic, although he invoked the monopoly power of large firms as an explanation for why such firms would be the drivers of technical change. This analysis suggests a similar empirical result but does not require that high-status organizations have monopoly power within their markets to gradually become the dominant forces in directing innovation. The relative contributions of status and monopoly power in shaping technical change could be observed if a technological domain could be studied from a period close to its inception. Biotechnology is one area that may afford the opportunity for such an exploration.

A second possible direction for future research would be an exploration of the connection between the technological ties that emerge among organizations in the elaboration of the technological network and the diverse array of interorganizational linkages. We expect that an organization's location in the technological network should both partially explain and be partially explained by these interorganizational relations.

Regardless of the direction of future research, the role-based ecology provides a sociological view of the rate and direction of technical change. We believe that these processes have yet to receive enough systematic research in sociology, given their importance in general and in particular their unique role in shaping the social organization of many economic activities.

## REFERENCES

- Albert, M. B., D. Avery, F. Narin, and P. McAllister. 1991. "Direct Validation of Citation Counts as Indicators of Industrially Important Patents." *Research Policy* 20:251-59.
- Anders, George. 1981. "Experts Say IBM's Entry Will Buoy Already Booming Software Industry." *Wall Street Journal* (August 19).
- Arthur, W. Brian. 1988. "Competing Technologies: An Overview." In *Technical Change and Economic Theory*, edited by Giovanni Dosi, Christopher Freeman, Richard Nelson, Gerald Silverberg, and Luc Soete. London: Pinter.
- Barley, Stephen R. 1990. "The Alignment of Technology and Structure." *Administrative Science Quarterly* 35:61-103.
- Barnett, William P., and Glenn R. Carroll. 1987. "Competition and Mutualism among Early Telephone Companies." *Administrative Science Quarterly* 32: 400-421.

- Basberg, Bjorn L. 1987. "Patents and the Measurement of Technological Change." *Research Policy* 16:131-41.
- Brodsky, Marc H. 1990. "Progress in Gallium Arsenide Semiconductors." *Scientific American* (February), pp. 68-75.
- Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- DiMaggio, Paul. 1986. "Structural Analysis of Organizational Fields: A Blockmodel Approach." Pp. 335-65 in *Research in Organizational Behavior*, edited by B. Staw and L. Cummings. Greenwich, Conn.: JAI Press.
- Dosi, Giovanni. 1984. *Technical Change and Industrial Transformation*. New York: St. Martin's.
- Dosi, Giovanni, and Luigi Orsenigo. 1988. "Coordination and Transformation: An Overview of Structures, Behaviors and Change in Evolutionary Environments." Pp. 13-37 in *Technical Change and Economic Theory*, edited by Giovanni Dosi, Christopher Freeman, Richard Nelson, Gerald Silverberg, and Luc Soete. London: Pinter.
- Durkheim, Émile. 1933. *The Division of Labor in Society*. New York: Free Press.
- Eerden, C. van der, and F. H. Saelens. 1991. "The Use of Science and Technology Indicators in Strategic Planning." *Long Range Planning* 24 (3): 18-25.
- Electronic News*. 1982-91. Selected issues.
- Farrell, J., and Garth Saloner. 1985. "Standardization, Compatibility and Innovation." *Rand Journal of Economics* 16:70-83.
- Gilder, George. 1989. *Microcosm*. New York: Simon & Schuster.
- Granovetter, Mark S. 1974. *Getting a Job*. Cambridge, Mass.: Harvard University Press.
- Green, Kenneth. 1992. "Creating Demand for Biotechnology: Shaping Technology and Markets." Pp. 164-84 in *Technological Change and Company Strategies: Economic and Sociological Perspectives*, edited by R. Coombs, P. Saviotti, and V. Walsh. London: Academic Press.
- Hannan, Michael T., and John Freeman. 1989. *Organizational Ecology*. Cambridge, Mass.: Harvard University Press.
- Hawley, Amos H. 1950. *Human Ecology*. New York: Ronald Press.
- Heckman, James J., and George J. Borjas. 1980. "Does Unemployment Cause Future Unemployment? Definitions, Questions, and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica* 47:247-83.
- Helm, Leslie. 1992. "U.S. Japan Battle of the Patents." *Los Angeles Times* (april 24).
- Hughes, Thomas P. 1983. *Networks of Power: Electrification in Western Society, 1880-1930*. Baltimore: Johns Hopkins University Press.
- . 1987. "The Evolution of Large Technological Systems." Pp. 51-82 in *The Social Construction of Technological Systems*, edited by Wiebe E. Bijker, Thomas P. Hughes, and Trevor J. Pinch. Cambridge, Mass.: MIT Press.
- Kalbfleisch, John D., and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Katz, Michael L., and Carl Shapiro. 1984. "Network Externalities, Competition, and Compatibility." *American Economic Review* 75:424-40.
- Ladd, David, David E. Leibowitz, and Bruce G. Joseph. 1986. *Protection for Semiconductor Chip Masks in the United States*. Weinheim. Verlag Chemie.
- Latour, Bruno. 1987. *Science in Action*. Cambridge, Mass.: Harvard University Press.
- Law, J., and Callon, Michel. 1988. "Engineering and Sociology in a Military Aircraft Project: A Network Analysis of Technological Change." *Social Problems* 35 (3): 284-90.
- Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney G. Winter.

1987. "Appropriating the Returns from Industrial Research and Development." *Brookings Papers on Economic Activity*. Washington, D.C.: Brookings Institution.
- March, James G. 1988 *Decisions in Organizations*. New York: Basil Blackwell.
- Marx, Karl. 1954. *Capital*. Moscow: Foreign Languages Publishing House.
- McClellan, William J., ed. 1981-91 *Status: A Report on the Integrated Circuit Industry*. Scottsdale, Ariz.: Integrated Circuit Engineering Corporation.
- McPherson, J. Miller. 1983. "An Ecology of Affiliation." *American Sociological Review* 48:519-32.
- Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159:56-63.
- Narin, Francis, Elliot Noma, and Ross Perry. 1987. "Patents as Indicators of Corporate Technological Strength." *Research Policy* 16:143-55.
- Nelson, Richard R., and Sidney G. Winter. 1982 *An Evolutionary Theory of Economic Change*. Cambridge, Mass.: Belknap Press.
- Office of Technology Assessment and Forecast, U.S. Department of Commerce, Patent, and Trademark Office. 1976. *Technology Assessment and Forecast*, 6th ed. Washington, D.C.: Government Printing Office.
- Orenstein, Susan. 1992. "Japanese Companies Shift Gears, Fight in Court." *Legal Times* (June 22).
- Perrow, Charles. 1967. "A Framework for the Comparative Analysis of Organizations." *American Sociological Review* 32:194-208.
- Pinch, Trevor J., and Wiebe E. Bijker. 1987. "The Social Construction of Facts and Artifacts." Pp. 17-50 in *The Social Construction of Technological Systems*, edited by Wiebe E. Bijker, Thomas P. Hughes, and Trevor J. Pinch. Cambridge, Mass.: MIT Press.
- Podolny, Joel M. 1993. "A Status-Based Model of Market Competition." *American Journal of Sociology* 98:829-72.
- . 1994. "Market Uncertainty and the Social Character of Economic Exchange." *Administrative Science Quarterly* 39 (3): 458-83.
- Ristelhueber, Robert. 1993. "Time for a Reality Check." *Electronic Business* (June): 99.
- Rohwer, Götz. 1993. *Transition Data Analysis*, ver. 5.2. Institut für Empirische und Angewandte Soziologie, Universität Bremen.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Schelling, Thomas. 1960. *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.
- Scherer, F. M. 1984. *Innovation and Growth: Schumpeterian Perspectives*. Cambridge, Mass.: MIT Press.
- Schumpeter, Joseph A. 1950. *Capitalism, Socialism and Democracy*, 3d ed. New York: Harper & Row.
- Staudenmaier, John M. 1985. *Technology's Storytellers: Reweaving the Human Fabric*. Cambridge, Mass.: MIT Press.
- Stinchcombe, Arthur L. 1990. *Information and Organization*. Berkeley and Los Angeles: University of California Press.
- Thompson, James D., and Frederick L. Bates. 1957. "Technology, Organization, and Administration." *Administrative Science Quarterly* 2:325-42.
- Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Information." *Rand Journal of Economics* 21 (1): 172-87.
- Tushman, Michael L., and Philip Anderson. 1986. "Technological Discontinuities and Organizational Environments." *Administrative Science Quarterly* 31:439-65.
- Tushman, Michael L., and Richard R. Nelson. 1990. "Introduction: Technology, Organizations, and Innovation." *Administrative Science Quarterly* 35:1-8.
- White, Harrison C. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 87:517-47.

Wilson, Robert W., Peter K. Ashton, and Thomas P. Egan. 1980. *Innovation, Competition, and Government Policy in the Semiconductor Industry*. Toronto: Lexington Books.

Woodward, Joan. 1958. *Management and Technology*. London: H.M.S.O.

———. 1965. *Industrial Organization: Theory and Practice*. New York: Oxford University Press.

# SYMPOSIUM ON APPLIED REGRESSION

## Statistical Methods for Comparing Regression Coefficients between Models<sup>1</sup>

Clifford C. Clogg  
*Pennsylvania State University*

Eva Petkova  
*Columbia University*

Adamantios Haritou  
*University of Macedonia*

Statistical methods are developed for comparing regression coefficients *between* models in the setting where one of the models is nested in the other. Comparisons of this kind are of interest whenever two explanations of a given phenomenon are specified as linear models. In this case, researchers should ask whether the coefficients associated with a given set of predictors change in a significant way when other predictors or covariates are added as controls. Simple calculations based on quantities provided by routines for regression analysis can be used to obtain the standard errors and other statistics that are required. Results are also given for the class of generalized linear models (e.g., logistic regression, log-linear models, etc.). We recommend fundamental change in strategies for model comparisons in social research as well as modifications in the presentation of results from regression or regression-type models.

### I. INTRODUCTION

The most common method used in empirical social research to compare two explanations of some specified dependent variable can be described as follows. A conventional explanation is represented by a linear model

<sup>1</sup> This research was supported in part by grant SBR-9310101 from the National Science Foundation and in part by the Population Research Institute, Pennsylvania State University, with core support from the National Institute of Child Health and Human Development (grant 1-HD28263-01). The authors are indebted to C. Himes for providing the data and some of the calculations used in table 5 and to L. A. Goodman, J. Hagenars, and four referees for helpful comments. Address correspondence to Clifford C. Clogg, Department of Sociology, Pennsylvania State University, University Park, Pennsylvania 16802.

that includes the set of predictors regarded as important factors in that explanation, as specified from the relevant theory and prior research. Suppose these are denoted as  $X = (X_1, \dots, X_p)$ . The alternative explanation is specified by adding to the original model a different set of predictors, say  $Z = (Z_1, \dots, Z_q)$ , giving a linear model that includes both  $X$  and  $Z$ . Both models are then compared in terms of variability explained, the relative magnitudes of the regression coefficients in the two models, and the magnitudes of the coefficients relative to their standard errors in either or both models. Such comparisons are usually made in order to determine whether the original explanation, the alternative explanation, or some combination of the two is most consistent with the data. Model comparisons of this general kind are ubiquitous in social research, as well as in other areas, and they are especially prominent whenever regression techniques are applied for the analysis of nonexperimental data, such as survey data. In this article, we recommend fundamental changes in the logic of model comparisons in social research.

Model comparisons formulated in the above manner generally involve two related but different questions. The first is whether the increment in prediction or explained variability obtained by adding  $Z$  is significant or important. Standard procedures that test the significance of the incremental contribution of  $Z$  given  $X$  can be used to answer this question. For linear models, the usual  $F$ -test that compares the more comprehensive model that includes  $X$  and  $Z$  and the less comprehensive model that includes only  $X$  can be applied. For other members of the class of generalized linear models, such as logistic regression, analogous chi-square statistics can be used. Tests such as these for incremental contributions are based on the assumption that the more comprehensive (or "full") model is the model that generated the data (see, e.g., Blalock 1979; Hanushek and Jackson 1977; Neter, Wasserman, and Kutner 1989). That is, in conducting tests for *incremental* variability, the logic for this type of model comparison is predicated on the assumption that the more comprehensive model is true or is conditional on the truth of this model.

A second question is equally important when regression techniques are used to compare explanations. Are the coefficients that describe the relationship between  $Y$  and  $X$  in the first model different from the coefficients that describe the relationship between  $Y$  and  $X$  in the second model where  $Z$  has been added and has hence been "statistically controlled"? Or is the regression relationship between  $Y$  and  $X$  *stable* across specifications? This question cannot be answered in a valid way with conventional methods used for the assessment of incremental improvement in prediction. Those procedures have no *direct* relationship to inferential questions pertaining to the comparison of regression coefficients between models,

except in some special cases. In short, the question of incremental improvement in prediction is different from the question of stability in regression relationships, and methods suited for the former question need not be valid for considering the latter question. In spite of this, researchers often compare the coefficients associated with  $X$  between the models by examining whether one or more of the regression coefficients associated with  $X$  is significantly different from zero in either model; or whether the level of significance of those coefficients differs between the models; or whether the apparent size of those coefficients differs between the two models. Comparisons of this general kind are invalid because (a) they ignore the fact that the coefficients in the first model are not independent of the coefficients in the second model and (b) they ignore the fact that both models cannot be "true" at the same time, except in the case where the two models are equivalent (i.e., when the  $Z$  factors do not contribute to the regression). The latter difficulty implies that the inferential logic used to make statistical assessments of this kind is simply not appropriate.

In this article we give simple procedures that can be used to compare regression coefficients across models when linear models or generalized linear models are used. We suggest modifications in the reporting of regression results that provide direct evidence about the relative plausibility of the two explanations. In our view, these methods ought to be used in practically all empirical research work where linear or generalized linear models are used to compare explanations. These methods utilize quantities that are routinely provided by standard computer programs for regression analysis and thus can be implemented with little extra work (and no new computer programs).

In the next two sections we cover the simplest case involving three variables. This case is covered first to fix concepts and to illustrate the logical and inferential issues that arise. It also gives an alternative framework for the analysis of the case that so often serves as a prototype for regression analysis in social research. For this case it is straightforward to compare a simple regression coefficient to a partial regression coefficient and thereby *test* whether "controlling" for a variable suppresses or enhances the relationship between two variables (see Bollen 1989, chaps. 2, 3; Duncan 1975).<sup>2</sup> In Section IV we give the generaliza-

<sup>2</sup> We hasten to add that the methods presented here assume only that a "single-equation" regression model is available, having a full and a reduced version. This is very different from a "multiple-equation" model, which is featured throughout the literature on structural equation models. The distinction between the two approaches should be clear in the next section; the fact that we use only a single-equation model, with full and reduced versions, should be clear throughout.



tion for multiple regression, where either or both  $X$  and  $Z$  stand for sets of variables. Results in this section can be applied in all cases where multiple regression is used. Section V gives results for the class of generalized linear models, and includes two examples. Examples throughout the article are designed to illustrate the logic of the suggested procedures and also to indicate how the calculations can be done using quantities routinely provided by software for regression analysis. Section VI summarizes recommendations for applying these methods and suggests several generalizations of the approach useful for different research settings.

## II. THE THREE-VARIABLE CASE

It is easiest to illustrate the logic of model comparison using the three-variable case where linear regression and least squares methods would be used. Let  $Y$ ,  $X$ , and  $Z$  denote the three variables:  $Y$  is the specified outcome variable of interest and  $X$  and  $Z$  are predictors. Let  $Y_i$ ,  $X_i$ ,  $Z_i$  denote the values that these variables take on for the  $i$ th case,  $i = 1, \dots, n$ , where  $n$  is the sample size. Two models are considered:

$$H_R: Y_i = \alpha_1 + \beta_1 X_i + \epsilon_i, \quad i = 1, \dots, n, \quad (1)$$

and

$$H_F: Y_i = \alpha_2 + \beta_2 X_i + \gamma Z_i + v_i \quad i = 1, \dots, n. \quad (2)$$

We assume that error terms are well behaved, at least in the "full" model,  $H_F$ . For example, we assume that  $v_i$  is  $N(0, \sigma_v^2)$  (normally distributed, homoscedastic errors) and that errors are not correlated with each other (or errors are independent of each other with the normality assumption). We let  $\sigma_\epsilon^2$  denote the error variance under the first model,  $\sigma_v^2$  the error variance under the second model; and we denote the standard estimators or estimates of these as  $\hat{\sigma}_\epsilon^2$ ,  $\hat{\sigma}_v^2$  (i.e., the mean-squared errors for the respective models).

Now let  $x_i = X_i - \bar{X}$ , the centered values, and let  $y_i$ ,  $z_i$  denote the centered values of the other two variables. The least squares estimator of  $\beta_1$  is

$$b_{yx} = \hat{\beta}_1 = \sum x_i y_i / \sum x_i^2 = s_{yx} / s_x^2 = r_{yx} s_y / s_x, \quad (3)$$

where  $s_x^2 = \sum x_i^2 / n$ ,  $s_{yx} = \sum y_i x_i / n$ ,  $s_y^2 = \sum y_i^2 / n$ , and  $r_{yx}$  denotes the ordinary coefficient of correlation between  $Y$  and  $X$ . In this section only, we write the least squares estimator of the slope as  $b_{yx} = \hat{\beta}_1$ , using a popular notation that reminds us of the fact that the estimator (and the model) do not "control" for  $Z$  (or any other variable). *Assuming that the*

*simple regression model is true*, the sampling variance of this estimator is

$$V(b_{yx}|H_R) = \sigma_\epsilon^2/(ns_x^2). \quad (4)$$

Replacing  $\sigma_\epsilon^2$  by  $\hat{\sigma}_\epsilon^2$  and taking the square root produces the estimated standard error of  $b_{yx}$ , say  $s(b_{yx})$ .

Using similar notation, the least squares estimator of  $\beta_2$  in model  $H_F$  is

$$b_{yx \cdot z} = \hat{\beta}_2 = (r_{yx} - r_{zx}r_{yz})s_y/[(1 - r_{zx}^2)s_x], \quad (5)$$

where we use the popular notation  $b_{yx \cdot z}$  for the partial regression coefficient (Blalock 1979). We see from this expression that  $b_{yx} = b_{yx \cdot z}$  if  $r_{zx} = 0$ . Assuming that model  $H_F$  is true, the variance of this estimator is

$$V(b_{yx \cdot z}|H_F) = \sigma_y^2/[ns_x^2(1 - r_{zx}^2)] \quad (6)$$

(see, e.g., Hanushek and Jackson 1977, chap. 2). We note that both models cannot be true simultaneously, except in the special (and usually trivial) case where  $\gamma = 0$ . The first logical issue that has to be addressed is *which* model of the two being considered is "true," or which can be assumed to be true for inferential purposes, or which ought to be assumed to be true for logically consistent inferences (whether formal or informal).

Next consider the comparison of the simple and the partial regression coefficient given above. In the population, let  $\delta = \beta_1 - \beta_2$ , and in the sample let

$$d = \hat{\delta} = b_{yx} - b_{yx \cdot z}. \quad (7)$$

(or  $d = \hat{\delta} = \hat{\beta}_1 - \hat{\beta}_2$ ). Note that  $d$  is simply the arithmetic difference between the simple regression coefficient and the partial regression coefficient. The quantity  $d$  is an important quantity in elementary analyses of causal systems involving three variables. We do not here invoke the causal imagery of three-variable recursive models, however, which would require that the "causal ordering" of  $X$  and  $Z$  be known, that an additional regression equation with either  $X$  or  $Z$  as the outcome be specified, that the variables included as predictors in both regressions are uncorrelated with the disturbances in the two models, and so on (Duncan 1975). It is very important to note that the logic involved here posits one outcome variable ( $Y$ ); there are two competing models for this single outcome: the reduced model and the full model.

If both regression coefficients are positive and their difference is positive, then controlling for  $Z$  as in model  $H_F$  would be said to partially explain the relationship between  $Y$  and  $X$ . Of course, if  $\beta_2 = 0$  (i.e.,

$b_{yx}$ , is not significantly different from zero), then we normally say that controlling for  $Z$  explains the relationship between  $Y$  and  $X$ , in which case  $\delta = \beta_1$  (or  $d$  is approximately the same as  $b_{yx}$ ). Using (3) and (5), we obtain

$$d = [r_{zx}(r_{yz} - r_{yz}r_{zx})s_y]/[(1 - r_{zx}^2)s_x]. \quad (8)$$

Note that  $d$  is identically zero when (a)  $r_{zx} = 0$  ( $X$  and  $Z$  are uncorrelated) or (b) the coefficient for  $Z$  (i.e.,  $\hat{\gamma}$ ), which is proportional to  $(r_{yz} - r_{yz}r_{zx})$ , is zero ( $Z$  does not predict once  $X$  is included). These results ought to be intuitive, because if the two predictors are uncorrelated then the simple and the partial regression coefficients are identical, and because if  $Z$  does not predict at all in model  $H_F$  then this model is identical to  $H_R$  in the sample. Neither condition will hold true exactly when nonexperimental data are used; condition (a) will hold approximately if levels of  $X$  have been randomly assigned.<sup>3</sup>

### III. INFERENCE ABOUT THE SIZE OF $\delta$ IN THE THREE-VARIABLE CASE

The previous section summarized results from least squares analysis of the linear regression involving three variables. Questions of inference—such as hypothesis tests, confidence intervals, and so on—were not addressed. We now elaborate on these issues with an eye on how  $b_{yx}$  and  $b_{yx}$ , ought to be compared using the formulae in the preceding section as a guide.

To make inferences about the simple regression ( $H_R$ ), it is ordinarily assumed that this model is "true." The usual  $t$ -statistic ( $t = b_{yx}/s(b_{yx})$ ) is based on this assumption, for example. But if  $H_F$  is also considered, then it is normally assumed that this model is true for purposes of making inferences about the parameters of that model. *Both models cannot be true simultaneously*, however, in the common setting where one model represents one causal or structural explanation and the other model represents a different causal or structural explanation. Of course, the reduced model is related to the full model as  $H_R = H_F + (\gamma = 0)$ , which is the usual way to represent the fact that the reduced model is nested in the full model. Of the two models,  $H_F$  can always be assumed true;  $H_R$  is true when the full model is true *and* when  $\gamma = 0$ . This fact has important consequences for any comparison of the two models, not just the inference about  $d = b_{yx} - b_{yx}$ .

<sup>3</sup> For practical purposes, we can assume that  $d$  will be nonzero with probability one in all cases where nonexperimental data are used and the sample size is relatively large. In the same situations,  $r_{zx}$  and  $\hat{\gamma}$  will both be nonzero as well.

We now make the assumption that inferences should be conditional on the full model,  $H_F$ , holding true. In this case, we usually regard  $\delta$  as the *bias* in the estimator of  $\beta_{yx} = \beta_1$  when  $Z$  is an omitted variable in the former model;  $d$  is then the natural estimator of "omitted-variable bias" (Hanushek and Jackson 1977; Arminger 1995). We do not have to make a strong assumption that model  $H_F$  is true in some deeply philosophical sense, but we do have to make inferences conditional on *something* in order to maintain logical consistency. In our view, it is most natural to assume that statistical inferences should be conditional on the full model. This is consistent with virtually all research practice where full and reduced models are compared, and so this assumption should not be controversial. For example, this assumption is applied to test for the significance of the incremental variation explained by  $Z$ , which would be assessed in this case with a  $t$ -statistic,  $t = \hat{\gamma}/s(\hat{\gamma})$ , or the square of this statistic (the  $F$ -statistic for incremental variability explained in this case). The same assumption is now used to draw inferences about the difference in regression coefficients.

Both  $b_{yx}$  and  $b_{yx \cdot z}$  are linear functions of the  $Y_i$  or the  $y_i$ . For example,

$$b_{yx} = \Sigma x_i y_i / (n s_x^2),$$

which shows that  $b_{yx} = \Sigma w_i^* y_i$ , with  $w_i^* = x_i / (n s_x^2)$ . The difference  $d = b_{yx} - b_{yx \cdot z}$  is also a linear function of the  $y_i$ ,

$$d = \Sigma w_i y_i. \quad (9)$$

Using (7), (3), and (5) and the definition of correlation leads to

$$w_i = x_i / (n s_x^2) - (s_x^2 x_i - s_{xz} z_i) / [n (s_x^2 s_z^2 - s_{xz}^2)], \quad (10)$$

where  $s_{xz} = \Sigma x_i z_i / n$ . The distribution of  $d$  is determined from this expression. For example, the variance of  $d$  is  $V(d) = \Sigma w_i^2 V(y_i)$ , assuming that observations are uncorrelated.<sup>4</sup>

If we assume that model  $H_F$  is true and hence condition on  $X$  and  $Z$ , then this means that we seek  $V(d|H_F) = V(d|X, Z)$ . This implies that we use the variance of  $Y$  given  $X$  and  $Z$  (or given model  $H_F$ ), which we write as  $V(y_i|H_F)$ . Now  $V(y_i|H_F) = V(y_i|x_i, z_i) = V(v_i) = \sigma_v^2$ , the error variance under the second model. Direct algebra gives  $\Sigma w_i^2 = r_{xz}^2 / [n s_x^2 (1 - r_{xz}^2)]$ , so the standard deviation of the difference  $d$  is

$$\sigma(d) = \sigma_v |r_{xz}| / [n s_x^2 (1 - r_{xz}^2)]^{1/2}. \quad (11)$$

<sup>4</sup> The term  $V(y_i)$  actually stands for the conditional variance of  $y_i$  given the predictors used (or some other condition), and so the assumption is that, given those predictors (or some other conditions), the observations are conditionally uncorrelated or conditionally independent.

The estimated standard error is obtained by replacing  $\sigma_v$  with  $\hat{\sigma}_v$ , the square root of the estimate of error variance in the full model. We see that this standard error is zero if  $r_{xz} = 0$ , in which case  $d = 0$  (simple and partial coefficients are identical). Because  $r_{xz} \neq 0$  can be assumed when nonexperimental data are analyzed and the sample is relatively large, the fact that  $\sigma(d) = 0$  when  $r_{xz} = 0$  does not affect the validity of the procedures given here. Note that we have assumed  $r_{xz}^2 < 1$  throughout (i.e.,  $X$  and  $Z$  are not colinear). Because we have conditioned on the full model (i.e., on both  $X$  and  $Z$ ), this correlation is not a parameter, although the magnitude of this correlation affects the standard error and hence the power of tests based on it.

Some additional formulae shed light on the logic of comparing coefficients between models in this three-variable case. For model  $H_F$ , the standard deviation of  $b_{yx \cdot z}$  is  $\sigma_v/[ns_x^2(1 - r_{xz}^2)]^{1/2}$ , and we see from (11) that

$$\sigma(d) = |r_{xz}| \sigma(b_{yx \cdot z}). \quad (12)$$

This expression shows the fallacy of making comparisons between the two coefficients assuming that the coefficients are independent or uncorrelated. If  $X$  and  $Z$  are correlated with  $r_{xz} = .5$ , for example,  $\sigma(d)$  is one-half of  $\sigma(b_{yx \cdot z})$ . The sample analogue to (12) is  $s(d) = |r_{xz}| s(b_{yx \cdot z})$ , using the ordinary standard error of the partial regression coefficient as calculated in the least squares analysis of the full model.

Assuming that  $H_F$  is true,  $d/s(d)$  follows a  $t$  distribution on  $n - 3$  *df* under the hypothesis that  $\delta = 0$ , where  $s(d)$  is obtained from (12) by substituting  $\hat{\sigma}_v$  for  $\sigma_v$ . It is easy to show also that

$$d = r_{xz} \hat{\gamma}_z / s_z$$

(see Clogg, Shihadeh, and Petkova 1992, p. 57), and from this it follows that  $t = d/s(d) = \hat{\gamma}_z / s(\hat{\gamma}_z)$ . In other words, the *t*-test for a significant difference between  $b_{yx}$  and  $b_{yx \cdot z}$  is equivalent to the test for the significance of  $\hat{\gamma}$  in model  $H_F$ , a result that holds *only* in the three-variable case considered here.

In a situation where this three-variable setup represents a realistic comparison of two explanations, we recommend calculating the difference  $d = b_{yx} - b_{yx \cdot z}$  and the standard error of this difference,  $s(d)$ , as above. In this case only, the significance of  $d$  can be determined from the significance of the coefficient of  $Z$ , or the "test for omitted variable bias" and the test for incremental explained variability are identical. In this case, we recommend replacing the usual variance estimator (or the standard error) for the quantity  $b_{yx}$  by an alternative estimator that recognizes that model  $H_F$ , and not  $H_R$ , is assumed to be true. We now give simple results that show how an adjustment of this kind can be made.

We seek  $V(b_{yx}|H_F)$ , not  $V(b_{yx}|H_R)$ , which is estimated under the simple regression model. Because  $b_{yx} = \sum w_i^* y_i$ , with  $w_i^* = x_i/(\sum x_i^2)$ , and  $V(y_i|H_F) = \sigma_v^2$ , we obtain the preferred estimator of the standard error,

$$s^*(b_{yx}) = \hat{\sigma}_v/(\sum x_i^2)^{1/2}, \quad (13)$$

taking note of the fact that  $V(b_{yx}|H_F) = \sigma_v^2 \sum w_i^{*2}$ . Let  $s(b_{yx})$  denote the standard error calculated for model  $H_R$ . Then we see that the corrected standard error is

$$s^*(b_{yx}) = s(b_{yx}) \hat{\sigma}_v / \hat{\sigma}_e.$$

Because  $V(U - W) = V(U) + V(W) - 2\text{cov}(U, W)$  for arbitrary random variables  $U$  and  $W$ , the covariance between  $b_{yx}$  and  $b_{yx \cdot z}$  can be obtained easily. That is,  $V(d) = V(b_{yx}) + V(b_{yx \cdot z}) - 2\text{cov}(b_{yx}, b_{yx \cdot z})$ , where the calculation assumes "variance under  $H_F$ " and "covariance under  $H_F$ ." The first three quantities have already been obtained. Using (12), (13), and (6), we find  $\text{cov}(b_{yx}, b_{yx \cdot z}) = \sigma_v^2/(\sum x_i^2)$ , that is,

$$\text{cov}(b_{yx}, b_{yx \cdot z}) = V(b_{yx}|H_F). \quad (14)$$

Finally, the correlation between the two regression coefficients is  $(1 - r_{xz}^2)^{1/2}$ , which is independent of the error variances, and this gives yet another way to see how closely associated the two estimators are.<sup>5</sup>

We conclude this section with another formula for the variance of  $d$  that turns out to be suitable for the general case. Using results given immediately above,  $V(d) = V[(b_{yx} - b_{yx \cdot z})|H_F]$  can be written as

$$V(d) = V(b_{yx \cdot z}) - V(b_{yx}|H_R) \sigma_v^2 / \sigma_e^2. \quad (15)$$

This means that the variance of  $d$  can be estimated as  $s^2(d) = s^2(b_{yx \cdot z}) - s^2(b_{yx}) \hat{\sigma}_v^2 / \hat{\sigma}_e^2$ , where these quantities are the squared standard errors and error variance estimates produced as standard output by a regression package. The square root of this quantity is  $s(d)$ .

### A. An Example

We illustrate these relationships with an example drawn from the 1990 General Social Survey (GSS). The dependent variable ( $Y$ ) is occupational prestige as calibrated on the 1980 census occupational classification system. Predictor  $X$  is respondent's education in years (EDUC) while the predictor  $Z$  is father's education (PAEDUC), both scored as years of

<sup>5</sup> Note that the covariance between the two coefficients is positive and hence the correlation is nonnegative, ranging over the interval  $[0, 1)$ . That is, the value of the simple regression coefficient and the value of the partial regression coefficient are positively correlated if  $r_{xz} \neq 0$ .

TABLE 1

COMPARISON OF COEFFICIENTS IN A THREE-VARIABLE CASE:  
OCCUPATIONAL PRESTIGE REGRESSED ON EDUCATION AND FATHER'S EDUCATION

Predictor	Model $H_R$	Model $H_F$	$d$	$t = d/s(d)$
Intercept .....	12.946 (2.912)	13.315 (2.977)	-.370 (.611)	-.605
EDUC .....	2.292 (.210)	2.351 (.232)	-.059 (.098)	-.605
PAEDUC .....		-.107 (.177)		
$\hat{\sigma}$ .....	11.964	11.974		
$R^2$ .....	.245	.246		
$n$ .....	368	368		

SOURCE.—GSS (1990).

NOTE.—Cases pertain to all males ages 20–64 with nonmissing codes for all variables. Note that  $\hat{y}/s(\hat{y}) = -.107/.177 = -.605$ . SEs are in parentheses. Calculations were carried out to more decimal places than are reported.

schooling completed. The sample is all male respondents ages 20–64 with nonmissing codes on all variables; both regressions were estimated on the same cases, which is necessary for application of the usual model comparisons as well as for the ones suggested here.<sup>6</sup>

Results appear in table 1. The coefficient associated with EDUC in the simple regression is  $b_{yx} = 2.292$ ; for the regression involving PAEDUC ( $Z$ ) the corresponding coefficient is  $b_{yz} = 2.351$ . Hence,  $d = -.059$  and  $s(d) = .098 = (.232^2 - .210^2 \hat{\sigma}_y^2 / \sigma_z^2)^{1/2}$  using (15) with substitution of the estimated error variances. The ratio of the two variance estimates is  $143.3821/143.1338 = 1.0017$ , which is greater than unity because the PAEDUC variable is unimportant relative to the loss of a degree of freedom. Even in this case where most analysts would delete the additional predictor (i.e., choose the reduced model), the estimated standard error of  $d$  is informative. Note that  $t = d/s(d) = \hat{y}/s(\hat{y}) = -.605$  (when more decimals are used). All of the preceding formulae can be checked out with this case. For example, (12) gives  $s(d) = .0980 = .4224 \times .2320 = r_{xz}s(b_{yz})$ . (Note that EDUC and PAEDUC have a Pearson correlation of .4224 in the sample of cases used for this regression). Note also that the difference in the intercepts can be studied in exactly the same way, for reasons that will be made clear in the next section. The ratio of the difference in intercepts to the standard error is

<sup>6</sup> A difficulty in applying these formulas to published regression results is that missing-case exclusions are often applied differently from model to model, with the result that different samples are used for each model.

also equal to the  $t$ -ratio for the added covariate ( $-.605$ ). If the full model with EDUC and PAEDUC were preferable, we could adjust the standard errors for the parameters in the reduced model by the ratio of the two estimates of standard deviations of the error terms (see [13] and surrounding text).

As shown earlier, the variance (and the standard error) of  $d$  was calculated assuming that the full model was true. And in this case, the test of the significance of  $d$  turns out to be equal to the test of the significance of  $\hat{\gamma}$  in the full model. Under standard assumptions, including normality of the errors  $v$ , in the full model, the ratio  $t = d/s(d) = \hat{\gamma}/s(\hat{\gamma})$  follows a  $t$  distribution exactly if the null hypothesis is true ( $\gamma = \delta = 0$ , in this case).

## B. Comparison with a Related Procedure

The above procedure is related to several large-sample tests for "model misspecification" as summarized in Godfrey (1988). Hausman's (1978) procedure is the most popular test of this kind. It will be considered briefly in relation to our procedure to illustrate important differences in logic and assumptions. Hausman's test applies to a wide range of model comparisons, not just to the type of model comparison considered in this article (we are not aware of examples in which the Hausman procedure is applied to the class of problems considered here, however). It has a very simple form for the problem of comparing two linear regression models where one is nested in the other.

The Hausman test in this setting gives  $V(d) = V(b_{yx\cdot}) - V(b_{yx})$ , which is closely related to our procedure (see [15]). The sample estimator of this variance is used to construct tests similar to those presented here. A different method is used, however, to estimate this variance. Often in *practice*, each variance on the right-hand side is calculated under the model used to obtain each quantity; that is,  $V(b_{yx\cdot})$  is estimated from model  $H_F$  and  $V(b_{yx})$  is estimated from model  $H_R$ . But the logic of the Hausman test is based on estimating the variance of  $d$  *under the reduced model*, so  $V(d)$  should be understood as  $V(d|H_R)$ . In other words, Hausman's test assumes that the reduced model is true, whereas we assume that the full model is true. Large-sample approximations are applied to estimate this variance in general settings where the method is applied. (In general settings where the Hausman method is applied, it can be difficult to determine the variance of parameter values in a full model assuming that the reduced model is true, and perhaps this accounts for the common practice of using the estimated variances of those parameters in each model.)<sup>7</sup>

<sup>7</sup> If Hausman's test is applied correctly in this case, in the sense that  $V(d|H_R)$  is actually calculated under the reduced model, then this test will be more powerful than our test if the reduced model  $H_R$  is true. If the full model is true, however, then



Hausman's procedure as applied to our setting consists of the following logic: (a) assume that the reduced model is true so that  $b_{yx}$  is both a consistent and an efficient estimator of  $\beta_{yx} = \beta_1$ ; (b) under this condition,  $b_{yx}$  is a consistent estimator of  $\beta_{yx} = \beta_1$  also; (c) if these assumptions are true, then  $V(d)$ —and tests of the significance of  $d$ —can be based on the above formula, with or without modification. There are many variations of the Hausman test in the literature. For example, Bentler and Chou (1993) present several statistics for assessing model improvement in the context of covariance-structure models that are closely related to Hausman's test and to the related Rao score tests. Their procedures are also based on an assumption that a given reduced model (or a model with one or more restrictions) is true. And their procedures have a different purpose, that of diagnosing lack of fit or of modifying the reduced model. These procedures are different from those presented here, because we assume that the full model is true.

Applying the Hausman method in the usual way gives the variance estimate,  $.2302^2 - .2101^2 = .0097$ , with square root (SE) .098, which is the same (to three decimal places) as the result derived from our method. But this result obtained because the ratio of mean-squared errors was approximately one in this case. If the variance of the difference is estimated without modification—without adjusting for the mean-squared errors—then in almost all cases where there is significant incremental variability due to  $Z$  the variance estimate derived from this approach will be *negative* (i.e., in cases where  $Z$  contributes significantly, the estimate of error variance will be reduced and the estimate of the variance  $V[b_{yx}]$  will be smaller than the estimate of the variance  $V[b_{yx}]$  as a consequence).<sup>8</sup> The similarity in this case between the Hausman procedure, as typically applied, and our procedure is due solely to the fact that  $Z$  does not contribute significantly to the regression, as reflected in the fact that the ratio of mean-squared errors is nearly unity. The Hausman test could not be used, without modification, in cases where  $Z$  contributes significantly and the standard error of  $b_{yx}$  is smaller than the standard error of  $b_{yx}$ . The procedure based on Hausman's framework is very different from the one we have proposed because it actually conditions on the reduced model holding true, whereas we have conditioned on the full model holding true. If the modification described above is not used in the application of the Hausman test, the estimator of the variance (or SE) of  $d$  can be negative. This difficulty does not arise with the

---

a test should be based on the full model, and the precision of the  $d$  value will be estimated in a valid way with our approach in this case.

<sup>8</sup> An example of this is that the Hausman procedure could not be applied without modification in the example using multiple regression in table 2 below.

method presented here (see, e.g., eq. [12]). Also, under the maintained assumptions (the full model is true, the errors have constant variance, etc.), the test of the significance of  $d$  is exact for small as well as large samples, whereas the Hausman procedure is a large-sample test with asymptotic justification.<sup>9</sup>

The conditions assumed in the Hausman test and in related misspecification tests of this general kind are very different from the conditions assumed here. Similar comments apply to the procedures covered next.

#### IV. MULTIPLE REGRESSION

We now give the generalization to multiple regression. The reduced model is represented as

$$H_R: Y_i = \alpha^* + \beta_1^* X_{i1} + \dots + \beta_p^* X_{ip} + \epsilon_i, \quad (16)$$

and the full model is represented as

$$H_F: Y_i = \alpha + \beta_1 X_{i1} + \dots + \beta_p X_{ip} \\ + \gamma_1 Z_{i1} + \dots + \gamma_q Z_{iq} + \nu_i, \quad (17)$$

where  $i = 1, \dots, n$ , and the same sample is used to evaluate both models.<sup>10</sup> There are  $p$  predictors in  $X$  and  $q$  predictors in  $Z$ ; the previous sections covered the case where  $p = q = 1$ . Denote the least squares estimators of the regression coefficients associated with  $X_k$  as  $b_k^*$  in the former (reduced) model and as  $b_k$  in the latter (full) model,  $k = 1, \dots, p$ . Call the corresponding vectors of regression coefficients  $b^*$  and  $b$ , respectively.<sup>11</sup>

Now let  $\delta_k = b_k^* - \beta_k$  and  $d_k = b_k^* - b_k$ , quantities that follow from earlier considerations of the three-variable case. In vector terms, we have

<sup>9</sup> If the errors  $\nu_i$  in the full model are normally distributed with constant variance, then  $d$  has an exact  $t$  distribution, and in large samples  $d$  will be normally distributed without the normality assumption.

<sup>10</sup> Note that we use an asterisk—instead of numbered subscripts as in the previous case—to denote the coefficient from the reduced model. Below we use the symbol  $b^*$  (or  $b_k^*$ ) for the (least squares) estimator in the reduced model, the symbol  $b$  (or  $b_k$ ) for the (least squares) estimator in the full model. We use the notation  $\beta$  to denote maximum-likelihood estimators. Of course, for linear models with normally distributed errors the least squares estimators are maximum-likelihood estimators.

<sup>11</sup> Note that the intercept can be regarded as the coefficient of the “predictor” defined as “the vector of ones.” Then  $X$  would have  $p + 1$  predictors, and we can examine differences in intercepts in the same way that we examine differences in the other coefficients. Inferences of this kind for intercepts could be of interest in the context of regression standardization (see, e.g., Clogg and Eliason [1986] and references cited therein).

$\delta = \beta^* - \beta$ ,  $d = b^* - b$ . In appendix A we show that  $V[(b^* - b)|H_F] = V(b) - V(b^*|H_F)$ . Note that  $V(b^*|H_F) = V(b^*|H_R)\sigma_v^2/\sigma_\epsilon^2$ , so that the variance-covariance matrix from the reduced model is multiplied by the ratio of error variances, as before, to produce the variance-covariance matrix under the full model. From these facts we obtain the simple expression,

$$s^2(d_k) = s^2(b_k) - s^2(b_k^*)\hat{\sigma}_v^2/\hat{\sigma}_\epsilon^2, \quad (18)$$

where  $s^2(\cdot)$  refers to the estimated variance of the quantity involved; and  $s^2(b_k^*)$  is the sampling variance (squared standard error) of the coefficient in the reduced model, as estimated in that model.<sup>12</sup> (Note that  $s^2(b_k^*) \times [\hat{\sigma}_v^2/\hat{\sigma}_\epsilon^2]$  is the estimate or estimator of  $V[b_k^*|H_F]$ .) The scale factor is simply the ratio of the two mean-squared errors. The square root of  $s^2(d_k)$  is the standard error of interest; we see that these can be calculated easily from standard output from the two regressions.

#### A. An Example Using Multiple Regression

To illustrate how these methods can be used, we consider regression models estimated with the 1990 GSS, with  $Y$  denoting occupational prestige and six predictors:  $X_1$  and  $X_2$ , spline-coded variables for years of education less than or equal to 12 or greater than 12, respectively;  $X_3$ , father's occupational prestige score (when the respondent was 16 years old);  $X_4$ , age in years;  $X_5$ , father's education in years;  $X_6$ , mother's education in years; thus  $p = 6$  predictors. The predictors added to form  $H_F$  are  $Z_1$ , a dummy variable for marital status (1 = currently married);  $Z_2$ , number of children in respondent's family;  $Z_3$ , unemployment history (dummy variable with 1 = ever unemployed);  $Z_4$ , a dummy variable indicating satisfaction with current residence; thus  $q = 4$  additional predictors. Results appear in table 2. The  $F$ -ratio for the comparison of the full and the reduced models is 2.837, which is significant at less than the .05 level with a reference distribution of  $F_{4,230}$ . The ratio of mean-squared errors is  $\hat{\sigma}_v^2/\hat{\sigma}_\epsilon^2 = 124.2156/128.1167 = .9696$ , and this quantity is used in (18) and could be used to correct the estimates of standard errors in the reduced model.

Even though adding  $Z$  contributes to explained variability in a modest way (an increase in  $R^2$  of .032), significant differences between the coefficients of several of the predictors in  $X$  are observed. The coefficient for post-high-school education (spline) decreases from 1.817 to 1.548 ( $d =$

<sup>12</sup> Compare (18) with (15). Further insights about this expression can be found in the previous section and in the appendix

TABLE 2

A COMPARISON OF TWO REGRESSION MODELS PREDICTING OCCUPATIONAL PRESTIGE

Predictor	Model $H_R$	Model $H_F$	$d_k = b_k^* - b_k$	$t = d_k/s(d_k)$
Intercept . . . . .	18.194 (8.273)	21.766 (8.508)	-3.573 (2.457)	-1.45
$X_1$ : education $\leq$ 12 years . . . . .	1.057 (.607)	1.176 (.605)	-.120 (.096)	-1.24
$X_2$ : education $>$ 12 years . . . . .	1.817 (.757)	1.548 (.754)	.269 (.119)	2.27
$X_3$ : father's occupation . . . . .	.126 (.072)	.106 (.072)	.019 (.009)	2.20
$X_4$ : age . . . . .	.139 (.067)	.072 (.072)	.067 (.028)	2.40
$X_5$ : father's education . . . . .	-.252 (.240)	-.206 (.239)	-.046 (.036)	-1.29
$X_6$ : mother's education . . . . .	-.019 (.034)	-.028 (.034)	.009 (.003)	2.64
$Z_1$ : marital status . . . . .		-3.393 (1.692)		
$Z_2$ : no. of children . . . . .		.090 (.561)		
$Z_3$ : unemployment history . . . . .		-2.803 (1.585)		
$Z_4$ : satisfaction with residence . . . . .		1.236 (1.514)		
$\sigma$ . . . . .	11.319	11.145		
$R^2$ . . . . .	.320	.352		

SOURCE.—GSS (1990)

NOTE.—See text for definition of variables. For both models  $N = 241$ .

.269), which is a difference of more than two standard errors. By controlling for the factors in  $Z$ , the coefficients for father's occupation, age, and mother's education have also changed in important ways. As a block, the coefficients for the six predictors have changed significantly as well. Because  $p = 6 \geq q = 4$ , the test for significance of the vector  $d = b^* - b$  is equivalent to the incremental  $F$ -statistic ( $F = 2.837$ , which is significant at less than the .05 level compared to the reference distribution of  $F_{4,230}$ ; see Clogg, Petkova, and Shihadeh 1992).<sup>13</sup>

<sup>13</sup> Note that the special case where  $p = q = 1$  was covered in the section on inference in the three-variable case. The squared  $t$ -statistic for the significance of  $d$  is identical to the incremental  $F$ -statistic for the significance of the predictor  $Z$ . When  $p < q$ , a related but different test for block differences has to be used. This is described briefly in the appendix and in Clogg et al. (1992). Finally, note that standard methods can be used to make simultaneous tests for some, but not all, differences. For example, if  $s$  differences were of interest, a simple (conservative) method of adjustment is to use  $\alpha^* = \alpha/s$  instead of  $\alpha$  for each of the  $s$  comparisons, which is an application of the Bonferroni method.

To summarize, we find that adding the four predictors in  $Z$  changes the relationship between  $Y$  and  $X$  significantly, and four of the coefficients have been singled out as the source of the difference. Finally, note that the tests based on the  $d_k$  values are very different from tests on the individual regression coefficients. For example, the coefficient for mother's education does not appear to be significantly different from zero in either model; however, the difference between these two coefficients is very significant. The explanation offered by the reduced model is thus quite different from the explanation offered by the full model, not only in terms of additional predictors used but also in terms of several coefficients describing the  $Y$ - $X$  relationship. Whether these differences are substantively important or not would depend on a variety of factors, but a rigorous and logically consistent assessment of these differences can be based on the  $d_k$ , the  $s(d_k)$ , and their ratios.<sup>14</sup>

#### B. A Note on Comparing Regression Coefficients between Groups

It is instructive to consider briefly the relationship between the proposed framework and methods for comparing regression relationships between groups. Suppose that the identical regression model corresponding to, say,  $H_R$ , were estimated in two groups such as males and females. The separate regressions can be retrieved from a regression model for both groups combined by adding a dummy variable  $D$  that denotes the group variable as well as the interaction of  $D$  with each of the predictors. If the error variance is the same in each group, then group differences in regression coefficients can be assessed in a valid way by testing the interaction terms, singly or jointly. When the error variances are not the same, different methods have to be used (see Gujarati 1988).

In large samples, the significance of the difference between the coefficient  $\hat{\beta}_m$  and the coefficient  $\hat{\beta}_f$  (the coefficient in the male sample and the coefficient in the female sample, say) can be assessed with the statistic

$$z = (\hat{\beta}_m - \hat{\beta}_f) / [s^2(\hat{\beta}_m) + s^2(\hat{\beta}_f)]^{1/2},$$

which follows a standard unit normal under the null hypothesis of equality of the two coefficients. The standard error of the difference is the square root of the sum of the two squared standard errors, assuming that the samples are independent.

Our procedures appear at first sight to be similar to the standard

<sup>14</sup> Significance of the incremental contribution of  $Z$  need not be associated with significant differences in the  $d_k$  values, even if the components of  $X$  and the components of  $Z$  are moderately correlated. An example demonstrating this appears in Clogg et al. (1992); also see the first example in the next section.

method outlined above, but actually they are very different. First, in our method the comparison of the simple regression coefficient (or any coefficient in a reduced model) and the partial regression coefficient (or any partial regression coefficient for predictors common to both the reduced and the full model) is based on the same sample, not different (independent) samples. The two quantities are correlated, as shown above. We might randomly split the sample into two groups and estimate model  $H_R$  for one sample and model  $H_F$  for the other, however, and then use a modification of the standard procedure. If the two samples are based on random selection, then we could, say, compare  $b_{yx}$  from the first sample with  $b_{yx}$  in the second sample. But the two models condition on different things, and this has to be taken into account. That is, a  $z$ -statistic analogous to that given above for the comparison of coefficients *from the same models* applied to independent samples (with  $b_{yx}$  used in place of  $\hat{\beta}_m$  and  $b_{yx}$  used in place of  $\hat{\beta}_f$ ) cannot be used to compare the simple and the partial regression coefficients, even if each is estimated from randomly selected subsamples. For example, the sampling variance of  $b_{yx}$  in the first sample needs to be modified to take account of the fact that the comparison should be conditional on the full model. An approximate way of doing this is to take  $s_{\star}^2(b_{yx}) = s^2(b_{yx}) \times MSE_F^*/MSE_R^*$ , where  $MSE_F^*$  denotes the estimate of error variance for the full model as estimated with the randomly selected subsample and  $MSE_R^*$  is the estimate of error variance for the reduced model as estimated with the other randomly selected subsample. Although this would be a valid procedure in large samples, it is not efficient, as can be seen from the fact that the full sample is not used to estimate the error variance under the full model. Similar comments would apply to the extensions of our method given next if they were compared to common methods for group comparisons of regression coefficients.<sup>15</sup>

## V. GENERALIZED LINEAR MODELS

The above procedures for linear regression will now be extended to the class of generalized linear models (McCullagh and Nelder 1989). That is, procedures are now developed for log-linear models for categorical data (Agresti 1990; Goodman 1978, 1984), logistic regression for binary outcome variables, most logit-type models for polytomous outcome variables, Poisson-regression models for event-count data, and most of the event history regression models used widely in social research. A frame-

<sup>15</sup> We are indebted to a referee for suggesting consideration of this approach.

work for analysis of this class of models was presented in Clogg, Petkova, and Shihadeh (1992); see Petkova and Clogg (1993) for derivations that are summarized here. We consider the general case as well as some specific examples.

The reduced model will be represented as

$$H_R: \eta_i = \alpha^* + \beta_1^* X_{i1} + \dots + \beta_p^* X_{ip}, \quad (19)$$

and the full model will be represented as

$$H_F: \eta_i = \alpha + \beta_1 X_{i1} + \dots + \beta_p X_{ip} + \gamma_1 Z_{i1} + \dots + \gamma_q Z_{iq}, \quad (20)$$

where  $i = 1, \dots, n$ . Here,  $\eta_i$  refers to a function of the expectation of  $Y$ , say  $\eta_i = g(\mu_i)$ , with  $E(Y_i) = \mu_i$ . The classical regression model in the previous sections is obtained by taking  $g$  as the identity "link." If  $Y_i$  refers to a count in cell  $i$  of a contingency table, then  $g$  is the logarithm and (19) and (20) represent two log-linear models for a contingency table with  $n$  cells. If  $Y$  is a binary dependent variable with  $E(Y_i) = \pi_i$ , the probability of "success," then taking  $g(\pi_i) = \log[\pi_i/(1 - \pi_i)]$  gives the standard logit model. If  $Y$  is multivariate, as with multinomial logistic regression,  $g$  is then a multivariate link function, usually a set of logits with one category chosen as a base. (If all variables are discrete so that the data can be arranged in a contingency table format, then the log link can be used. The log-frequency model can be reformulated in the usual way as a model for logits; see Goodman 1978.) The framework given here also applies to multivariate outcomes with appropriate modifications. We thus deal with a class of models that is wide enough to cover most applications of regression methods in social research.<sup>16</sup>

For model  $H_R$ , denote the "weight matrix" associated with a given generalized linear model as  $W^*$ . Denote the corresponding weight matrix for model  $H_F$  as  $W$ . For the ordinary log-linear analysis of contingency tables, with  $F_i$  denoting the expected frequency in cell  $i$ ,  $W^*$  is the diagonal matrix with the expected frequencies under model  $H_R$  on the diagonal, and  $W$  is the diagonal matrix with the expected frequencies under model  $H_F$  on the diagonal. For logistic regression involving a binary dependent variable, the matrices are also diagonal matrices, with the

<sup>16</sup> The conditions leading to  $\beta^* = \beta$  ("Z is collapsible") for the class of generalized linear models are different from those in the ordinary linear model. Agresti (1990) summarizes these. The main restriction in our formulation is that we consider only members of this class of models that do not include a special overdispersion parameter for unmodeled heterogeneity.

entry in the  $i$ th diagonal defined as  $w_i = \pi_i(1 - \pi_i)$  (for the full model), with a corresponding entry for the reduced model (see Agresti 1990, chap. 3). These weight matrices are readily available with standard routines or can be easily calculated from those routines.<sup>17</sup>

Now let  $\beta$  and  $\beta^*$  denote the relevant coefficient vectors where the intercept has been included, both then with  $p + 1$  elements. The variance of  $\beta^*$  under model  $H_R$  is the  $(p + 1) \times (p + 1)$  matrix,

$$V^*(\beta^*) = (X^T W^* X)^{-1}, \quad (21)$$

where the asterisk denotes the reduced model. This quantity is used to calculate the standard errors reported for this model, and this matrix is available from almost all software packages that would be used to estimate this model. The variance of  $\beta^*$  under model  $H_F$  is the so-called sandwich estimator,

$$V(\beta^*) = (X^T W^* X)^{-1} (X^T W X) (X^T W^* X)^{-1} \quad (22)$$

(see, e.g., White [1980] and Godfrey [1988] for the general form of this estimator). Finally, the covariance between  $\beta^*$  and  $\beta$  can be shown to be equal to

$$\text{cov}(\beta^*, \beta) = V^*(\beta^*) = (X^T W^* X)^{-1}. \quad (23)$$

Derivations appear in Petkova and Clogg (1993).<sup>18</sup>

The variance of the estimated difference ( $\delta = \beta^* - \beta$ ) is thus the  $p \times p$  matrix,

$$V(\delta) = V(\beta) + V(\beta^*) - 2\text{cov}(\beta^*, \beta), \quad (24)$$

where  $V(\beta)$  is the variance matrix of the regression coefficients for  $X$  in the full model (standard output), and the remaining two quantities are defined in (22) and (23). Note that the variance and the covariance terms in this expression are calculated under the full model, which is consistent with the inferential logic put forth earlier for the linear model. For com-

<sup>17</sup> These weight matrices are used to define the variance estimators, based on the information function, and so are used implicitly to calculate standard errors in routines using maximum likelihood estimation. The upper-left-hand block of the information matrix is  $-(X^T W X)$  for the full model, e.g.; and the estimated information is obtained by substituting  $\hat{W}$  in this expression.

<sup>18</sup> Note that the covariance under the full model is (asymptotically) equal to the variance of  $\beta^*$  under the *reduced* model, which is different from the corresponding *exact* variance obtained from the ordinary linear model. In this section all variance and covariance calculations are "large-sample" quantities and must be understood in this light.



parisons of individual coefficients, the standard error is obtained by taking the square root of the relevant diagonal entry in this matrix. This result can be simplified further. Using (21)–(23), we obtain

$$V(\hat{\beta}) = V(\beta) + V^*(\beta^*)(X^T W X) V^*(\beta^*) - 2[V^*(\beta^*)], \quad (25)$$

where  $V(\beta)$  is the variance matrix for the coefficients associated with  $X$  in the full model and  $V^*(\beta^*)$  is the variance matrix for the coefficients associated with  $X$  in the reduced model. These quantities will be provided as standard output, perhaps with a user-specified option, in virtually all programs for the analysis of generalized linear models or special cases of them. The only additional quantity to be calculated which is not a by-product of the calculations is the matrix  $(X^T W X)$ . This matrix can be obtained by inverting the variance-covariance matrix of parameter estimates for the full model, taking the matrix block of the result that corresponds to elements of  $X$  and taking care to ensure numerical accuracy. This matrix can be obtained in other ways as well; see below.

It can be noted that the usual Hausman (1978) test, if applied to this particular comparison of nested models, would approximate the variance as  $V(\hat{\beta}) - V^*(\hat{\beta}^*)$ , using the estimated variance matrices under each model. Both the variance estimator and the suggested test proposed here differ from the Hausman procedure. Estimates obtained from Hausman's formula for the variance can be negative (or nonpositive definite; see also Holly 1988 and Sec. III above), whereas the above estimator is nonnegative. The inferential logic associated with our approach is related to but different from that associated with the Hausman procedure. We do not assume, for example, that the reduced model is true; instead we condition on the full model (or assume that it is true) just as in the standard comparison of the two models using the difference between the two-model chi-square statistics.

#### A. Example: Collapsibility in Contingency Tables

To illustrate how the above results can be used in the analysis of contingency tables, we analyze the classic Lazarsfeld (1948) panel data analyzed with log-linear models in Goodman (1973) and with other methods by others. While the example illustrates how to apply the techniques in general settings where contingency tables are used, we also develop in this context a method for analyzing the "two faces" of panel analysis (Duncan 1981).

In this example, there are four dichotomous items leading to a four-way cross-classification. By comparing full and reduced models as specified below, the hypothesis of collapsibility can be examined in a rigorous way. Collapsibility with respect to a given parameter or quantity means that

TABLE 3

LAZARSFELD PANEL DATA:  
OBSERVED FREQUENCIES AND FITTED FREQUENCIES UNDER TWO MODELS

CELL (A, B, C, D)	OBSERVED FREQUENCIES	FITTED FREQUENCIES	
		Model $H_F(\hat{P})$	Model $H_R(\hat{P}^*)$
(1,1,1,1) \ . . . . .	129	128.10	72.07
(1,1,1,2) . . . . .	3	5.45	13.70
(1,1,2,1) . . . . .	1	.80	7.61
(1,1,2,2) . . . . .	2	.65	41.62
(1,2,1,1) . . . . .	11	11.58	18.68
(1,2,1,2) . . . . .	23	20.87	3.55
(1,2,2,1) . . . . .	0	.07	1.97
(1,2,2,2) . . . . .	1	2.48	10.79
(2,1,1,1) . . . . .	1	1.99	12.81
(2,1,1,2) . . . . .	0	.08	2.44
(2,1,2,1) . . . . .	12	12.11	1.35
(2,1,2,2) . . . . .	11	9.81	7.40
(2,2,1,1) . . . . .	1	.33	38.44
(2,2,1,2) . . . . .	1	.60	7.31
(2,2,2,1) . . . . .	2	2.01	4.06
(2,2,2,2) . . . . .	68	69.06	22.20

SOURCES — Lazarsfeld (1948) and Goodman (1973).

NOTE — *A* and *C* refer to vote intention, Republican vs Democrat, at the first and second interviews, respectively, *B* and *D* refer to opinion of the Republican candidate, favorable vs unfavorable, at the first and second interviews, respectively. See text for description of the models  $N = 266$ .

the given parameter is the same in the *marginal* table as in the *full* table with no variables collapsed, i.e., that marginal and partial “association” are equivalent (see, e.g., Agresti [1990]; Clogg et al. [1992]; and references in the latter source). For the analysis of collapsibility involving dichotomous variables, relatively straightforward methods can be used (Ducharme and Lepage 1986). But the analysis of collapsibility using existing methods is not so straightforward when polytomous variables are involved, particularly when the table has more than three dimensions. In contrast, the methods given here can be applied to examine collapsibility for arbitrary contingency tables.

Table 3 gives the observed frequencies in the cross-classification of the four items as well as estimated expected frequencies under two models. Variables *A* and *B* refer to vote intention (1 = Republican, 2 = Democrat) and candidate opinion (1 = favorable toward Republican, 2 = unfavorable) in the first interview, and *C* and *D* are similarly defined variables for the second interview. Model  $H_F$  is the model fitting the marginals  $\{(AB), (AC), (BD), (CD)\}$ . This model fits the data well with a

likelihood-ratio chi-square value of 6.74 on 7 *df*.<sup>19</sup> We see that this model includes across-time relationships, or "lagged effects," for the two repeated measures. That is, this model includes the interaction between *A* and *C* and the interaction between *B* and *D*. But the interaction between opinion at time 1 and intention at time 2 and the interaction between intention at time 1 and opinion at time 2 are both deleted in this model (i.e., the *AD* and *BC* interactions are excluded). The model does, however, include the contemporaneous "partial" association of the two variables at time 2, in the *C*–*D* interaction term. The reduced model  $H_R$  is the model fitting the marginals  $\{(AB), (CD)\}$ , which says that the joint variable *AB* is independent of the joint variable *CD*. This "reduced" model fits poorly, with a likelihood-ratio chi-square value of 497.33, 12 *df*. We would not take this model seriously as a description of the distribution in the four-way table, but this model has a special property that permits the analysis of collapsibility. This property is exploited next.

Collapsing over the first-wave measurements, *A* and *B*, gives the  $2 \times 2$  cross-classification of *C* and *D*, the second-wave measurements. The frequencies in this table are

$$\begin{pmatrix} 142 & 27 \\ 15 & 82 \end{pmatrix}$$

(levels of *C* in the rows, levels of *D* in the columns). The logarithm of the odds ratio is  $\hat{\phi}^* = 3.3587$ , with standard error  $s(\hat{\phi}^*) = .3506$ . The log-linear model for this marginal table, with expected frequencies denoted by  $F_{++kl}$ , is

$$\log(F_{++kl}) = \lambda^* + \lambda_{C(k)}^* + \lambda_{D(l)}^* + \lambda_{CD(kl)}^*,$$

using the standard definition of parameters. We find (Agresti 1990; Goodman 1973) that  $\lambda_{CD(11)}^* = \phi^*/4$ ,  $\hat{\lambda}_{CD(11)}^* = \hat{\phi}^*/4$ ,  $s(\hat{\lambda}_{CD(11)}^*) = s(\hat{\phi}^*)/4$ . The asterisk refers to the fact that the parameter or quantity pertains to the marginal ( $C \times D$ ) table. Now because model  $H_R$  for the full table posits independence between *AB* and *CD*, the parameters, parameter estimates, and standard errors for the above model can be obtained from the parameters, parameter estimates, and standard errors, respectively, for the log-linear model for the full table corresponding to  $H_R$ , that is, from

$$H_R: \log(F_{ijkl}^*) = \lambda^* + \lambda_{A(i)}^* + \lambda_{B(j)}^* + \lambda_{AB(ij)}^* + \lambda_{C(k)}^* + \lambda_{D(l)}^* + \lambda_{CD(kl)}^*.$$

It is also the case that the marginal association between *A* and *B* can be obtained directly from the *A*–*B* interaction in this expression, and so on. We now examine rigorously whether the marginal *C*–*D* interaction given

<sup>19</sup> This model is equivalent to Goodman's (1973, table 5) model  $H_{10}$ .

from the collapsed table (or in the model  $H_R$ ) is different from the *partial*  $C$ - $D$  interaction in the model  $H_F$ .<sup>20</sup> The full model is

$$H_F: \log(F_{ijkl}) = \lambda + \lambda_{A(i)} + \lambda_{B(j)} + \lambda_{AB(ij)} + \lambda_{C(k)} \\ + \lambda_{D(l)} + \lambda_{CD(kl)} + \lambda_{AC(ik)} + \lambda_{BD(jl)},$$

where we have arranged terms to correspond to those used in  $H_R$  above. Except for notational differences, we now have a full and reduced model of the kind introduced at the beginning of this section.

Given the equivalence between the model  $H_R$  for the full table and the "model" giving parameters for marginal association in both the  $A$ - $B$  and the  $C$ - $D$  tables after collapsing, the marginal and partial association can be compared statistically using the same kind of tabular layout as was used earlier for the standard regression problem. That is, we can study rigorously the "two faces" of panel analysis (Duncan 1981) to determine if the contemporaneous marginal association at wave two (the relationship between  $C$  and  $D$  in the marginal  $C \times D$  table) is different from the partial association between  $C$  and  $D$  in the full table (full model) "controlling" for across-time relationships. The relevant quantities appear in table 4. All results were obtained using (21)-(24), with the diagonal weight matrices estimated from the corresponding fitted values in table 3.

We see from table 4 that the  $C$ - $D$  marginal interaction value, .840 corresponding to a log-odds ratio of 3.36 (odds ratio of 28.75), is not significantly different from the  $C$ - $D$  partial interaction value, .736 corresponding to a partial log-odds ratio of 2.95 (partial odds ratio of 19.03). The difference (.103) is not significant compared to the standard error (.076). (The standard error is obtained by taking the square root of the relevant entry in the variance matrix defined above.) In summary, "controlling" for across-time relationships included in model  $H_F$  does not change the inference about the relationship between intention and candidate opinion at the second wave in a significant way, so there does not need to be a special explanation for how contemporaneous association is produced by lagged "effects" (Duncan 1981). That is, in this example the two faces in the two-wave panel are indistinguishable. Similar comments

<sup>20</sup> Other quantities can be examined in this fashion, but it does not make sense to compare the marginal  $A$ - $B$  interaction to the "partial"  $A$ - $B$  interaction because variables  $C$  and  $D$  refer to second-wave measurements and thus should not be "controlled for" when studying the association between variables at the first wave. In  $H_R$  we have included the terms relevant for studying the  $AB$  marginal table to indicate how such a comparison could be made in a context where this type of comparison is relevant. Exactly the same results would be obtained if the  $A$ ,  $B$ , or  $AB$  terms were deleted from  $H_R$ , even though the fitted frequencies for the reduced model would change if we deleted those terms to form the reduced model.

TABLE 4

A COMPARISON OF TWO LOG-LINEAR MODELS FOR THE LAZARSFELD PANEL DATA:  
MARGINAL AND PARTIAL ASSOCIATION COMPARED

Parameter	Model $H_R$	Model $H_F$	$d_k = b_k^* - b_k$	$z = d_k/s(d_k)$
<i>C</i> .....	.284 (.088)	.081 (.201)	.203 (.181)	1.12
<i>D</i> .....	-.009 (.088)	-.095 (.114)	.104 (.073)	1.43
<i>CD</i> .....	.840 (.088)	.736 (.112)	.103 (.076)	1.35
<i>A</i> .....	.251 (.076)	.210 (.194)	NR	
<i>B</i> .....	.063 (.076)	.113 (.100)	NR	
<i>AB</i> .....	.612 (.076)	.152 (.108)	NR	
<i>BD</i> .....		.936 (.111)		
<i>AC</i> .....		1.721 (.198)		

NOTE.—Comparisons not relevant (NR) for *A*, *B*, and *A-B* terms, see text.

apply to the *C* and *D* main-effect terms, although we would not ordinarily be interested in these terms as such. (Because none of the three terms differs significantly between the models, we can say that the entire *CD* table "structure" is not affected by controlling for lagged effects.) Note that significant differences might have been expected because the full model is a dramatic improvement over the reduced model in this case.

Straightforward modifications of the above approach can be used to examine any collapsibility question if collapsibility is defined in terms of parameters in log-linear models. If polytomous variables were used, for example, the comparisons between two-factor or multifactor interactions across models formed in the above fashion would lead to chi-square tests with more than one degree of freedom (see, e.g., Clogg et al. 1992). Note that the logic is applicable to contingency-table modeling in general, such as for the analysis of mobility tables, and in even more general circumstances. The reduced model need not correspond to a collapsed table, but it must be nested within the full model in order to apply this framework. For the analysis of a given mobility table, for example, the reduced model might correspond to quasi-perfect mobility having a parameter for each cell on the main diagonal, and the full model might correspond to an association model that also includes parameters for the cells on the main diagonal. This approach could be used to determine

whether adding the parameters for off-diagonal association in the full model changes the inference about the magnitude of "persistence" effects for the cells on the main diagonal. The general suitability of this approach for answering questions of this kind is one advantage of this approach compared to others in the contingency-table area (see, e.g., Ducharme and Lepage 1986; Whittemore 1978). The approach is directly suited for conventional analyses of collapsibility where a *variable* is collapsed, as indicated by the above example.

### B. Logistic Regression: Example

Our final example pertains to logistic regression where categorical or continuous predictors are used. Note that the literature on collapsibility in contingency tables is not relevant when one or more of the predictors are continuous. (If predictors are categorical, the previous example can be used as a guide for the analysis, taking into account the standard relationship between a log-linear model for frequencies and the logit model.) The necessary formulae are still (21)–(25), and we simply need to supply the matrix  $(X^T W X)$ . This matrix can be calculated as follows. For the full model  $H_F$ , denote the fitted logit model as

$$\hat{\Phi}_i = \hat{\alpha} + \hat{\beta}X_i + \hat{\gamma}Z_i.$$

The predicted probability for the  $i$ th case is thus  $\hat{\pi}_i = \exp(\hat{\Phi}_i)/[1 + \exp(\hat{\Phi}_i)]$ , and the  $i$ th element of  $\hat{W}$  is hence  $\hat{w}_i = \hat{\pi}_i(1 - \hat{\pi}_i)$ . The matrix of interest can then be calculated in a variety of ways, for example, by transforming the variables and obtaining the sum of squares and cross-products for an OLS regression. Generalizations of this procedure can be used for any member of the class of generalized linear models; each is characterized by a different weight matrix, which depends on the link function.<sup>21</sup>

Table 5 presents results from two logistic regression models estimated from the Supplement on Aging (SOA) of the 1984 National Health Interview Survey (NHIS). The models predict whether the respondent reported "fair or poor health" ( $Y = 1$ ) versus "good health" ( $Y = 0$ ). The first model ( $H_R$ ) predicts the logit of  $\text{prob}(Y = 1)$  from self-reports of "symptomatic" conditions or "asymptomatic" conditions summarized in the notes to the table. Model  $H_F$  adds  $q = 7$  other predictors designed to measure demographic and other characteristics of the respondent, in-

<sup>21</sup> We can also obtain the estimate of the matrix  $X^T W X$  by inverting the variance-covariance matrix of parameters for the full model; the upper-left-hand block of this matrix corresponding to the predictors in  $X$ , including the intercept, is the desired quantity.

TABLE 5

## LOGISTIC REGRESSION MODELS PREDICTING SELF-REPORT OF POOR OR FAIR HEALTH

Predictor	Model $H_R$	Model $H_F$	$\delta_k$	$z = d_k/s(d_k)$
Constant .. . . .	-2 136 (.057) [.054]	- 747 (.192)	-1.379 (.182)	-7.57
Symptomatic . . . . .	1.628 (.060) [.056]	887 (.067)	742 (.020)	37.14
Asymptomatic . . . . .	.613 (.085) [.079]	508 (.092)	106 (.015)	7.22
Limitations . . . . .		432 (.009)		
Age . . . . .		- 018 (.003)		
Female . . . . .		-.412 (.045)		
Nonwhite . . . . .		546 (.071)		
Not married . . . . .		- 176 (.048)		
Education . . . . .		- 549 (.034)		
Metro/noncentral . . . .		- 040 (.054)		
Nonmetro . . . . .		148 (.148)		
Model $\chi^2$ . . . . .	1,125.66	4,677.24		

SOURCE.—Supplement on Aging, 1984 National Health Interview Survey.

NOTE.—The data pertain to  $N = 15,808$  individuals age 55 and older in 1984. Symptomatic conditions include rheumatic heart disease, coronary heart disease, angina pectoris, myocardial infarction, any other heart attack or stroke or cerebrovascular accident, broken hip, rheumatic fever, Alzheimer's disease, cancer of any kind ("ever had"), or arthritis or rheumatism, diabetes, aneurysm, blood clots ("had in past 12 months"). Asymptomatic conditions include hypertension, hardening of the arteries, osteoporosis ("ever had"), or varicose veins ("had in past 12 months"). "Limitations" is number of functional limitations, residence was measured in three categories: metropolitan-central-city (deleted category), metropolitan-not-central-city (metro/noncentral), and nonmetropolitan (nonmetro). SEs are in parentheses, SEs under the full model for the coefficients in the reduced model are in brackets.

cluding residence. Comparing the coefficients for symptomatic conditions and asymptomatic conditions between models indicates how these other status measures influence the relationship between reported conditions and self-reported health; the comparison thus can be used to quantify the degree to which self-reported conditions translate into perceptions of poor health. This relationship has implications for the survey measurement of health status (see Johnson and Wolinsky 1993).

Because the sample size is so large, it is not surprising that the observed differences for the two predictors of interest change significantly. It seems clear that controlling for the status and other characteristics ( $Z$  factors) produces a dramatic change in the coefficient for symptomatic conditions (the difference is .74), which is over 37 SEs from zero. The coefficient

for asymptomatic conditions changes less (the reduction is .11), but is still highly significant, over 7 SEs from zero. Controlling for the status, residence, and other factors in  $Z$  thus changes the relationship between reported health conditions and self-reported health in a dramatic way. This finding can either be construed as a problem in using (subjective) self-reports of health or as an indication of the extent to which health conditions are mediated by status and other factors and so translate objective health conditions into self-reported perceptions of health.

A block test of no significant differences in any of the three coefficients in the reduced model based on this approach leads to chi-squared statistic 1,794.35, 3 *df*. Obviously, most of the block difference is due to the dramatic change in the coefficient for symptomatic conditions. Note that all variance (or SE) calculations as well as this chi-square value condition on the full model being true. That the full model is an improvement over the reduced model cannot be doubted; the likelihood-ratio statistic comparing the two models is  $3,551.58 = 4,677.24 - 1,125.66$ , 8 *df*, using the "model chi-square" values given in table 5. In this case, we thus find both a significant increment due to adding  $Z$  and significant change in the coefficients associated with  $X$ . The previous example was a case where the increment by adding  $Z$  was very significant, but still there was not significant change in the main parameter of interest.

This procedure can also be contrasted with the Hausman (1978) test, as this procedure would be used for this type of model comparison. This procedure conditions on the reduced model, which is definitely not "true" given the above results. Then  $V(\hat{\delta})$  would normally be approximated as  $\hat{V}(\hat{\beta}) - \hat{V}^*(\hat{\beta}^*)$ , and the Hausman chi-square statistic is  $H = \hat{\delta}'[V(\hat{\delta})]^{-1}\hat{\delta}$ . The value of  $H$  in this case is 882.24, which is very different from the value obtained above.<sup>22</sup> Once again the inferential logic distinguishes this approach from others: as with ordinary tests for incremental contributions, our approach conditions on the full model holding true. Straightforward modifications of the procedure suggested here can be used to develop the estimators and tests for other regression-type models in the class of generalized linear models (Poisson regression for event-count data, multinomial logistic regression, many models for event history data, etc.).

## VI. CONCLUSION

Examples used to illustrate the application of this method of comparing regression coefficients were designed mostly to show how calculations

<sup>22</sup> Both statistics lead to rejection, but they are each specified under different conditions, so it is not a surprise that these values differ. Note that the reduced model would be decisively rejected and so a test assuming that the reduced model is true could lead to very different results compared to a test assuming that the full model is true.



can be done easily with standard output from routines for regression analysis. With the exception of the example comparing the two faces of panel analysis (in the previous section), we do not claim that these examples demonstrate well the power or the flexibility of the suggested framework. We thus conclude with some observations on the potential role of these methods in theoretical explanation and in answering some methodological questions that are central in several substantive areas.

The concept of full and reduced models used throughout this article is obviously more than a statistical framework that makes inferences consistent and valid. In sociology as well as other areas, different theoretical explanations or "causal structures" are most often represented in this fashion. If a given theory or causal structure is represented by our equivalent of the reduced model ( $H_R$ ) with specified predictors  $X$ , then it is important to examine the *stability* of the coefficients associated with  $X$  when various modifications are made in the explanation or causal structure. The most common way to modify a given explanation is to consider various other models that also include  $X$  but include additional factors ( $Z$ ). This logic is closely related to the discovery and validation of "laws" (Pratt and Schlaifer 1988), and is or ought to be a vital aspect of theoretical analysis where alternative explanations are truly pitted against each other (cf. Jasso 1988). Direct application of this framework would seem to confront this general issue and also give explicit "rules" for elaborating particular theories represented as linear equations. In the vast majority of applications of regression methods for either causal inference or theoretical explanation in social research, analysts are implicitly examining questions of stability or change in regression coefficients (or regression relationships) under alternative specifications. Our framework can be used to examine those questions explicitly, logically, and with valid statistical methods. We do not claim, however, that this framework gives an automatic recipe for causal inference, only that the comparisons developed here provide assistance in this task. (The problem of causal inference with regression models and "nonexperimental data" does not reduce to just stability in coefficients under alternative specifications; see Smith [1990] and references cited therein.)

Although not a panacea for causal inferences using regression methods, there are several important methodological or substantive questions that can be addressed by this general framework. We conclude by mentioning some of these. (1) In an experiment or a quasi-experiment where  $X$  denotes the treatment levels "assigned," comparing the coefficient of  $X$  in the reduced model with the coefficient of  $X$  in various full models with added covariates provides information about the reliability of the causal inference that would be drawn. If the difference  $d$  is substantial or significant under alternative full models, for example, then this indicates

that the randomization (or quasi-randomization) was not effective. (2) It is often the case that either  $X$  or  $Z$  is categorical or can be categorized in various ways. By specifying the categorizations as full and reduced versions of a model of interest, the loss of information or the change in the inference associated with different ways of combining categories can be examined rigorously by means of the framework examined here. An example in a contingency-table setting appears in Clogg and Shihadeh (1994, chap. 2). (3) Many questions related to contextual analysis or group effects (see Manski 1993; Bryk and Raudenbush 1992) can be illuminated by application of the suggested framework. How to aggregate the groups used to define social context is an obvious case in point. If, say,  $X$  denotes individual-level factors and  $Z$  denotes group-level factors, in the simplest version of an analysis of this kind, then contextual analysis ought to be telling us not just whether  $Z$  contributes incrementally but whether the individual-level inferences (coefficients for  $X$ ) change once the macrolevel factors are included. (4) Finally, at least some questions concerning selectivity bias can be addressed using this approach. Many models for selective samples have been proposed (cf. Manski [1993] and references cited therein). In many cases, selectivity is modeled by introducing the formal equivalent of added covariates ( $Z$ ). Estimating this model and finding that these factors add "significantly" to the regression is not the same as demonstrating that the inference without an adjustment for selection is wrong for every predictor. Sharper inferences about the role of selection (Which relationships are affected by the selection process?) could be obtained by applying this framework.

The above extensions have not been worked out in sufficient detail so that prescriptions can be made, although these seem to be natural areas in which to extend this approach. But application of the framework put forth here would strengthen the comparison of regression coefficients across or between models that characterizes so much empirical research in sociology and other areas of social research. The modifications suggested for presenting findings from regression models featured in every example in this article would better communicate the evidence for model-to-model differences in regression relationships that figure so prominently in contemporary research practice.

## APPENDIX

### Derivations for the Linear Model

The main result in Section IV, special cases of which appear in Section III, is that  $V(b^* - b) = V(b) - V(b^*)$ , where the variance is taken under the full model and the standard assumptions are applied (homoscedastic-

ity, normality of errors). This result does not appear in the literature, to our knowledge, so a complete derivation is provided here.<sup>23</sup>

We let  $X$  denote the predictors in the reduced model including the vector of ones for the intercept (so that  $X$  is of order  $n \times [p + 1]$ ). The variance of  $b^*$  under the full model is

$$V(b^*) = (X^T X)^{-1} \sigma_v^2, \quad (\text{A1})$$

where  $\sigma_v^2$  is the error variance in the full model, not the reduced model. This expression follows from the fact that the least squares estimator of  $b^*$  is  $(X^T X)^{-1} X^T Y$ . The variance of  $b$  (coefficient vector for  $X$  in the full model) is

$$V(b) = A \sigma_v^2, \quad (\text{A2})$$

where

$$A = [X^T X - X^T Z(Z^T Z)^{-1} Z^T X]^{-1}. \quad (\text{A3})$$

This matrix is derived in the usual way by partitioning the design matrix for the full model as  $Q = [X, Z]$ , so that  $Q^T Q$  is the partitioned matrix

$$\begin{pmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z \end{pmatrix}.$$

The matrix  $A$  in (A3), of order  $(p + 1) \times (p + 1)$ , is the entry in the inverse of this matrix corresponding to the upper-left position. This quantity times the error variance under the full model gives the variance formula in (A2).

By similar methods, it can be shown (Clogg et al. 1992, p. 59) that the difference in the coefficient vectors is  $d = b^* - b = MY$ , where

$$M = [(X^T X)^{-1} X^T - A X^T + A X^T Z(Z^T Z)^{-1} Z^T], \quad (\text{A4})$$

where  $A$  is defined in (A3). From this it follows that  $V(d) = MM^T \sigma_v^2$  (under the full model). We now show that  $MM^T = A - (X^T X)^{-1}$ , from which it follows that  $V(d) = V(b) - V(b^*)$ , again with the understanding that the variance is evaluated under the full model.

From the definition of  $A$  in (A3), we obtain the relationship,

$$X^T X - A^{-1} = X^T Z(Z^T Z)^{-1} Z^T X, \quad (\text{A5})$$

<sup>23</sup> Although we use results in Clogg et al. (1992), the simplifications given here were not included in this work.

which is used repeatedly below. Term by term multiplication of  $M$  and  $M^T$  using (A4) gives

$$\begin{aligned} MM^T &= (X^T X)^{-1} (X^T X) (X^T X)^{-1} - (X^T X)^{-1} X^T X A \\ &\quad + (X^T X)^{-1} X^T Z (Z^T Z)^{-1} Z^T X A - A X^T X (X^T X)^{-1} + A X^T X A \\ &\quad - A X^T Z (Z^T Z)^{-1} Z^T X A + A X^T Z (Z^T Z)^{-1} Z^T X (X^T X)^{-1} \\ &\quad - A X^T Z (Z^T Z)^{-1} Z^T X A + A X^T Z (Z^T Z)^{-1} (Z^T Z) (Z^T Z)^{-1} Z^T X A. \end{aligned}$$

Using (A5), this simplifies to

$$\begin{aligned} MM^T &= (X^T X)^{-1} - A + (X^T X)^{-1} [X^T X - A^{-1}] A - A \\ &\quad + A X^T X A - A [X^T X - A^{-1}] A \\ &\quad + A [X^T X - A^{-1}] (X^T X)^{-1} \\ &\quad - A [X^T X - A^{-1}] A + A [X^T X - A^{-1}] A. \end{aligned}$$

Canceling terms leads directly to

$$\begin{aligned} MM^T &= (X^T X)^{-1} - A + A - (X^T X)^{-1} \\ &\quad - A + A X^T X A - A X^T X A + A + A - (X^T X)^{-1}, \end{aligned}$$

which simplifies to

$$MM^T = A - (X^T X)^{-1},$$

the result to be shown. This means that

$$V(d) = A \sigma_v^2 - (X^T X)^{-1} \sigma_v^2, \quad (\text{A6})$$

which is  $V(b) - V(b^*)$ , with the variance calculated under the full model.

The variance is estimated by substituting  $\hat{\sigma}_v^2$ , the mean-squared error under the full model, with  $n - p - q - 1$  degrees of freedom. When the design matrix for the full model is of full rank, the matrix  $MM^T$  is positive definite, so negative variance estimates will not arise (see related comments in Sec. III). Because  $d$  is a linear combination of the observations ( $Y$ ),  $d$  is normally distributed with mean  $\delta$  and variance defined above; in small as well as large samples the exact distribution of the standardized value of any  $d_k$  value follows the  $t$  distribution, a result used in Sections III and IV (see Clogg et al. 1992 for details).

To test the hypothesis that a given  $\delta_k = \beta_k^* - \beta_k = 0$ , the ordinary  $t$ -statistic can be used. To test the hypothesis that the entire block is zero (i.e.,  $\delta = 0$ ), the statistic

$$F^* = d^T [\hat{V}(d)]^{-1} d / p^* \quad (\text{A7})$$

can be used, where  $p^* = \min(p + 1, q)$ . The divisor follows from the fact that the rank of  $M$  is  $p^*$ , and if  $q < p + 1$  then a generalized inverse must be used; however, in the latter case, the statistic in (A7) is equivalent to the usual  $F$ -statistic for the incremental contribution of  $Z$  given  $X$ . The statistic in (A7) follows the  $F_{p^*, n-p-q-1}$  distribution under the hypothesis of no differences. If a block is chosen that is a subset, say  $p' < p + 1$ , of the original elements of  $X$ , then  $p^* = \min(p', q)$ . Subsets of the coefficients can be examined in an analogous manner.

## REFERENCES

- Agresti, Alan. 1990. *Categorical Data Analysis*. New York: Wiley.
- Arminger, Gerhard. 1995. "Mean Structures." In *Handbook of Statistical Modeling in the Social Sciences*, edited by Gerhard Arminger, Clifford C. Clogg, and Michael Sobel. New York: Plenum.
- Bentler, Peter M., and Chih-Ping Chou. 1993. "Some New Covariance Structure Improvement Statistics." Pp. 235-255 in *Testing Structural Equation Models*, edited by Kenneth A. Bollen and J. Scott Long. Newbury Park, Calif.: Sage.
- Blalock, Hubert M. 1979. *Social Statistics*, 2d ed. rev. New York: McGraw-Hill.
- Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
- Bryk, Anthony S., and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage.
- Clogg, Clifford C., and Scott R. Eliason. 1986. "On Regression Standardization for Moments." *Sociological Methods and Research* 14:423-46.
- Clogg, Clifford C., Eva Petkova, and Edward S. Shihadeh. 1992. "Statistical Methods for Analyzing Collapsibility in Regression Models." *Journal of Educational Statistics* 17:51-74.
- Clogg, Clifford C., and Edward S. Shihadeh. 1994. *Statistical Models for Ordinal Variables*. Newbury Park, Calif.: Sage.
- Ducharme, G. R., and Y. Lepage. 1986. "Testing Collapsibility in Contingency Tables." *Journal of the Royal Statistical Society, ser. B*, 48:197-205.
- Duncan, Otis Dudley. 1975. *Introduction to Structural Equation Models*. New York: Academic Press.
- . 1981. "Two Faces of Panel Analysis: Parallels with Comparative Cross-Sectional Analysis and Time-Lagged Association." *Sociological Methodology* 21: 281-318.
- Godfrey, L. G. 1988. *Misspecification Tests in Econometrics*. Cambridge: Cambridge University Press.
- Goodman, Leo A. 1973. "Causal Analysis of Data from Panel Studies and Other Kinds of Surveys." *American Journal of Sociology* 78:1135-91.
- . 1978. *Analyzing Qualitative/Categorical Data*. Cambridge, Mass.: Abt Books.
- . 1984. *Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.
- Gujarati, Damodar N. 1988. *Basic Econometrics*, 2d ed. New York: McGraw-Hill.
- Hausman, Jerry A. 1978. "Specification Tests in Econometrics." *Econometrica* 46: 1251-71.
- Hanushek, Eric A., and John E. Jackson. 1977. *Statistical Methods for Social Scientists*. New York: Academic Press.

- Holly, Albert. 1988. "A Remark on Hausman's Specification Test." *Econometrica* 50:749-59.
- Jasso, Guillermina. 1988. "Principles of Theoretical Analysis." *Sociological Theory* 6:1-20.
- Johnson, R. J., and F. D. Wolinsky. 1993. "The Structure of Health Status among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health." *Journal of Health and Social Behavior* 34:105-21.
- Lazarsfeld, Paul F. 1948. "The Use of Panels in Social Research." *Proceedings of the American Philosophical Society* 92:405-10.
- Manski, Charles F. 1993. "Identification Problems in the Social Sciences." *Sociological Methodology* 23:1-56.
- McCullagh, Peter M., and John A. Nelder. 1989. *Generalized Linear Models*, 2d ed. London: Chapman & Hall.
- Neter, John, William Wasserman, and Michael H. Kutner. 1989. *Applied Linear Regression Models*. Homewood, Ill.: Irwin.
- Petkova, Eva, and Clifford C. Clogg. 1993. "Collapsibility and 'Omitted Variables Bias' in Generalized Linear Models." Manuscript. Columbia University, Department of Biostatistics.
- Pratt, J. W., and R. Schlaifer. 1988. "On the Interpretation and Observation of Laws." *Journal of Econometrics* 39:23-52.
- Smith, Herbert L. 1990. "Specification Problems in Experimental and Nonexperimental Social Research." *Sociological Methodology* 20:59-93.
- Whittemore, Alice S. 1978. "Collapsibility of Multidimensional Contingency Tables." *Journal of the Royal Statistical Society, ser. B*, 40:328-40.
- White, Halbert. 1980. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica* 48:817-38.

# The Impact of Random Predictors on Comparisons of Coefficients between Models: Comment on Clogg, Petkova, and Haritou<sup>1</sup>

Paul D. Allison  
*University of Pennsylvania*

As Clogg, Petkova, and Haritou (1995) correctly observe, regression analysts are often intensely concerned with what happens to the coefficient of a predictor variable when additional variables are introduced into a regression model. Unfortunately, this concern is typically expressed in comparisons that lack any measure of statistical reliability. To remedy this deficiency, Clogg, Petkova, and Haritou (hereafter CPH) propose a set of methods for testing whether the change in a regression coefficient (or set of coefficients) is statistically significant. These methods have the virtues of simplicity and applicability to a wide class of generalized linear regression models.

CPH deserve credit for identifying an important but overlooked problem, and their solutions are both clever and elegant. Nevertheless, I believe that their proposed methods suffer from a fundamental flaw: they make unrealistic assumptions about the sampling properties of the predictor variables. Not surprisingly, this strategy leads to a substantial simplification of their methodology. But in doing so, it exposes the analyst to the risk of highly misleading conclusions.

My main objective is to explain why I think their assumptions are problematic and to suggest what may go wrong as a consequence. I shall focus primarily on the three-variable linear model since it embodies all the critical issues without the distracting complications of the multivariable and nonlinear cases. And if the methods are deficient in the three-variable case, there is little point in generalizing them to more complicated situations. I shall also suggest some alternative methods for both three-variable and multivariable linear models.

<sup>1</sup> For helpful suggestions, I am indebted to Nicholas Christakis, Arthur Goldberger, Herbert Smith, and Richard Waterman. Direct correspondence to Paul D. Allison, Department of Sociology, University of Pennsylvania, 3718 Locust Walk, Philadelphia, Pennsylvania 19104-6299. E-mail: allison@ssdc.sas.upenn.edu

## THE THREE-VARIABLE CASE

As far as possible, I will utilize the notation of CPH. The basic situation is this: we first regress  $Y$  on  $X$ . Then we regress  $Y$  on  $X$  and  $Z$ . We want to know if there is a significant difference between the two coefficients for  $X$ . Like CPH, I assume that the data are generated by a "full" model:

$$Y_i = \alpha + \beta_{yx \cdot x} X_i + \beta_{yz \cdot x} Z_i + v_i, \quad i = 1, \dots, n, \quad (1)$$

where  $v_i$  satisfies the usual assumptions of the linear model.<sup>2</sup> All derivations are based on this model, not on the "reduced" model that excludes  $Z$ .

The basic aim is to test the null hypothesis that  $\delta = 0$ , where  $\delta = \beta_{yz} - \beta_{yz \cdot x}$ , and  $\beta_{yz}$  is the population least squares regression coefficient of  $Y$  on  $X$  alone. That is,  $\beta_{yz} = \sigma_{xy}/\sigma_x^2$ , where  $\sigma_{xy}$  is the covariance of  $X$  and  $Y$ , and  $\sigma_x^2$  is the variance of  $X$ . An unbiased estimator of  $\delta$  is just  $d = b_{yz} - b_{yz \cdot x}$  where the  $b$ 's are the sample least squares estimators of the  $\beta$ 's. It can be shown that

$$\delta = \beta_{yz \cdot x} \left( \frac{\sigma_{xz}}{\sigma_x^2} \right), \quad (2)$$

which is the population analogue of CPH's formula

$$d = b_{yz \cdot x} \left( \frac{s_{xz}}{s_x^2} \right), \quad (3)$$

where  $s_{xz}$  is the sample covariance between  $X$  and  $Z$  and  $s_x^2$  is the sample variance of  $X$ . We see then that  $\delta = 0$  if  $\beta_{yz \cdot x} = 0$  or if  $\sigma_{xz} = 0$ , that is, if either  $Z$  has no effect on  $Y$ , controlling for  $X$ , or if  $X$  and  $Z$  are uncorrelated. In other words, there are two different ways that the null hypothesis can be true. But CPH focus only on the first possibility. They conclude that an appropriate test statistic for this null hypothesis is just the usual  $t$ -statistic for the null hypothesis that  $\beta_{yz \cdot x} = 0$ , namely, the ratio of the least squares coefficient  $b_{yz \cdot x}$  to its estimated standard error. In this setting, then, one needs only to check the significance of the added variable. If its coefficient is significantly different from zero, we conclude that the coefficient for the initial variable has also *changed* significantly.

<sup>2</sup> Specifically, I assume that  $E(v_i | X_i, Z_i) = 0$ ,  $V(v_i | X_i, Z_i) = \sigma^2$ , and  $\text{cov}(v_i, v_j | X_i, Z_i, X_j, Z_j) = 0$  for all  $i$  and for all  $j \neq i$ . For exact inferences, one also needs normality of  $v_i$ .



This is a surprising result, one that CPH claim applies only in the three-variable case. What they do not say is that there is an analogous and even more startling result in the multivariable case: suppose that a single  $Z$  variable is added to a model with several  $X$  variables. Under CPH's procedure, the  $t$ -statistic for the *change* in each  $X$  coefficient is identical to the  $t$ -statistic for the  $Z$  coefficient. In other words, when a single variable is added to a regression model, *either every coefficient changes significantly or none do*. I suspect that most experienced regression analysts will be troubled by this property. When a variable is added to a regression equation, some coefficients may change greatly, others hardly at all. Is it plausible that a procedure for testing changes in coefficients should be insensitive to the magnitudes of the changes?

CPH reach these unusual conclusions because of the way they treat the predictor variables  $X$  and  $Z$ . In textbook treatments of regression analysis, it is common to assume that the predictor variables are fixed or nonstochastic. That means that, in repeated sampling, all the values of the predictors stay the same from sample to sample, and only the values of the dependent variable change. In fact, however, this assumption is only plausible for experimental designs or for stratified sampling when there is one stratum for every combination of values of the predictor variables. But the vast bulk of social science research is nonexperimental, and no one does stratified sampling in that way. So why do textbook writers make such a patently unrealistic assumption? Because it enormously simplifies the algebra and it is relatively innocuous. For inferences about a correctly specified model, you get the same (or equivalent) results whether you treat the predictor variables as random or fixed. For example, in a two-variable linear model, the standard result for fixed  $X$  is that  $V(b_{yx}) = \sigma_e^2 / ns_x^2$ . When  $X$  is random, we have  $V(b_{yx}) = \sigma_e^2 E(1/ns_x^2)$ . These results are operationally equivalent because, in the random case, we estimate the expectation by its sample value.

Results are not equivalent, however, when the aim is to compare a full model with a restricted model. When predictor variables are fixed, for example, it is well known that deleting  $Z$  from the model reduces the variance of the coefficient of  $X$ . But when  $X$  and  $Z$  are random, Binkley and Abbot (1987) showed that the variance of the coefficient of  $X$  may increase substantially when  $Z$  is deleted. Similarly, in comparing the performance of alternative variable selection methods, Breiman and Spector (1992) concluded: "There can be startling differences between the  $x$ -fixed and  $x$ -random situations."

Although CPH do *not* assume that the predictor variables are fixed, they do something equivalent: they make all their inferences *conditional*

on the realized values of the predictor variables.<sup>3</sup> Here's a crucial example: from CPH's formula (11), it is easily shown that the conditional variance of  $d$  is

$$V(d|X, Z) = V(b_{y \cdot x}) \left[ \frac{s_{xz}}{s_x^2} \right]^2. \quad (4)$$

Noting that  $s_{xz}/s_x^2$  is the sample regression coefficient of  $Z$  on  $X$ , we can also write this as

$$V(d|X, Z) = V(b_{y \cdot x}) b_{zx}^2. \quad (5)$$

This result is correct as far as it goes, but treating (5) as *the* variance of  $d$  assumes that  $s_{xz}$  and  $s_x^2$  are constants, not random variables that can vary from sample to sample. In the next section I show that when  $X$  and  $Z$  are random, the *unconditional* variance of  $d$  is larger than (5), possibly much larger.

#### ALTERNATIVE METHODS FOR THE THREE-VARIABLE CASE

In the appendix, I derive the unconditional variance of  $d$  in the multivariable case. Specializing to the three-variable case, we have

$$V(d) = V(b_{y \cdot x}) E(b_{zx}^2) + \beta_{y \cdot x}^2 V(b_{zx}). \quad (6)$$

Comparing (6) with (5), we see that the first component of the sum in (6) corresponds to (5) except that we have taken the expectation of  $b_{zx}^2$ . Since the second component of (6) is nonnegative, the unconditional variance exceeds the conditional variance except when  $\beta_{y \cdot x} = 0$ .

In practice, a consistent estimator for this variance is

$$s^2(d) = s^2(b_{y \cdot x}) b_{zx}^2 + b_{y \cdot x}^2 s^2(b_{zx}). \quad (7)$$

All of these quantities are available from standard computer output. Here,  $s(b_{y \cdot x})$  is just the estimated standard error of  $b_{y \cdot x}$ , and  $s(b_{zx})$  is the estimated standard error of  $b_{zx}$ . The usual formula for  $s(b_{zx})$  depends, of course, on the assumption that  $Z$  can be expressed as a linear function of  $X$  with an error term satisfying the standard conditions. If this is

<sup>3</sup> CPH appear to confuse conditioning on the *model* with conditioning on the predictor variables when they write  $V(d|H_F) = V(d|X, Z)$  where  $H_F$  refers to the full model. While they are correct that it is necessary to choose some model as "true" in order to calculate expectations and variances, that does not imply that one must also condition on the variables in that model. In statistical terminology, conditioning on a model has no technical meaning, while conditioning on a random variable (or variables) is a well-defined operation.

implausible (e.g., if  $Z$  is dichotomous), one can use robust variance estimates (White 1980) that are now available in many regression programs. Having calculated  $s^2(d)$ , one can construct the test statistic  $d/s(d)$ , which will have approximately a standard normal distribution under the null hypothesis that  $\delta = 0$ . Alternatively, an approximate 95% confidence interval can be calculated as  $d \pm 2s(d)$ .

For hypothesis testing, another solution is based on likelihood theory. Assuming a trivariate normal distribution for  $X$ ,  $Y$ , and  $Z$ , one can maximize the likelihood, subject to the constraint that  $\delta = 0$ , and construct a likelihood ratio test by comparing the constrained likelihood with the unconstrained likelihood. The chi-square statistic is calculated as twice the positive difference in the two log likelihoods. Notice, however, that we can carry out the constrained estimation in two distinct steps. First constrain  $\beta_{yx} = 0$ , and then constrain  $\sigma_{xz} = 0$ . Pick whichever likelihood is larger and contrast that with the unrestricted likelihood (which must be larger still). This amounts to doing two likelihood-ratio tests, one for each constraint, and using the chi-square that is smaller.

This logic could be extended to more conventional tests for regression coefficients. First regress  $Y$  on both  $X$  and  $Z$ , and do a standard test for  $\beta_{yx} = 0$ . Then regress  $Z$  on  $X$  (or, equivalently,  $X$  on  $Z$ ) and test for significance of that coefficient. Reject the null only if both tests lead to rejection. Note that under this procedure, the probability of rejecting the null hypothesis (and concluding that the coefficient of  $X$  has changed) is always lower than it is for the CPH test because the rejection region of my proposed test is a proper subset of their rejection region.

#### A MONTE CARLO STUDY

To get an empirical comparison of the conditional and unconditional methods, I generated random samples from a "population" in which  $X$  and  $Z$  had a bivariate normal distribution with means of zero, variances of one, and a correlation of  $\rho$ , where  $\rho$  was assigned values of .00, .30, and .50. I then generated  $Y$  according to the equation

$$Y = 1 + 2X + \gamma Z + E, \quad (8)$$

where  $E$ , an "unobserved" random disturbance, was normally distributed with a mean of zero, a variance of 100, and was independent of  $X$  and  $Z$ . The coefficient  $\gamma$  had values of either zero or three. When  $\gamma = 0$ , it follows that  $\delta = 0$ . When  $\gamma = 3$ , we have  $\delta = 3\rho$ . For each of the six combinations of  $\rho$  and  $\gamma$ , I drew 500 random samples of size  $n = 100$ . In each sample, I calculated  $d$ , the conditional and unconditional estimates of the variance of  $d$ , and nominal 95% confidence intervals

TABLE 1

COMPARISON OF CONDITIONAL AND UNCONDITIONAL METHODS WITH SIMULATED DATA,  $\gamma = 3$ 

$\rho$	$\delta$	$V(d)$	CONDITIONAL		UNCONDITIONAL	
			Mean $\hat{V}(d)$	% Covered	Mean $\hat{V}(d)$	% Covered
.00	.00	.102	.014	15	.116	99
.10	.30	.115	.021	51	.124	94
.30	.90	.203	.114	80	.211	94
.50	1.50	.429	.360	92	.439	96

around  $\delta$  using both the conditional and unconditional variance estimates.

Tables 1 and 2 give the results. I will first examine table 1, in which  $\gamma = 3$ . The third column gives the actual variance of  $d$  across the 500 samples, which is a model-free estimate of the true variance. The fourth column reports the mean of the estimated variances of  $d$  using CPH's conditional estimator. When  $\rho = 0$ , the actual variance is more than seven times the average estimated conditional variance, a truly horrendous performance. As  $\rho$  gets larger, the CPH estimator does somewhat better, although even at  $\rho = .50$  it is still only 84% of the actual value.

These variance underestimates are reflected in the percentages of confidence intervals that actually include the true value, shown under "% covered." For  $\rho = 0$ , only 15% of the CPH confidence intervals include the true value  $\delta = 0$ . Equivalently, a two-tailed test of the null hypothesis (which is true in this case) would be rejected in 85% of the samples. For  $\rho = .10$ , the coverage improves substantially to 51%, and for  $\rho = .50$  it is almost at the nominal level. By contrast, the unconditional methods perform very well. Although the average of the variance estimates is always greater than the actual variance, the overestimate is never greater than 14%. And the coverage of the confidence intervals is fairly close to the nominal value of 95%.

Now turn to table 2, in which  $\gamma = 0$ . From equation (6), we see that when the true coefficient of  $Z$  is zero, the unconditional variance reduces to the conditional variance. Hence, there is no reason in this case to expect the unconditional method to do any better than the conditional method, and the results in table 2 bear that out. In fact, the unconditional estimates do somewhat worse. While the results for the conditional methods are nearly ideal, the unconditional variance estimates are too high, and the confidence intervals are too large. These biases are greatest when  $\rho = 0$  and become trivial when  $\rho = .50$ . Even in the worst case, how-

TABLE 2  
COMPARISON OF CONDITIONAL AND UNCONDITIONAL METHODS WITH SIMULATED  
DATA,  $\gamma = 0$

$\rho$	$\delta$	$V(d)$	CONDITIONAL		UNCONDITIONAL	
			Mean $\hat{V}(d)$	% Covered	Mean $\hat{V}(d)$	% Covered
.00	.00	.009	.010	95	.022	100
.10	.00	.021	.020	93	.032	100
.30	.00	.108	.110	95	.120	99
.50	.00	.367	.365	95	.375	97

ever, the unconditional confidence intervals are conservative in that the probability of rejecting the true null hypothesis is *less* than the nominal  $\alpha$  level.

One should not take these simulations as definitive, however. Only a few sets of parameter values and one sample size were considered. The assumption that the predictor variables are normally distributed is also unduly restrictive. Nevertheless, I think the results are sufficiently clear to raise substantial doubts about the appropriateness of CPH's procedures for the kinds of data that are typically analyzed by social scientists.

#### THE MULTIVARIABLE CASE

The generalization to the multivariable case involves a few new issues. In matrix notation the true model is

$$Y = X\beta + Z\gamma + \nu, \quad (9)$$

where  $Y$  and  $\nu$  are  $n \times 1$  vectors,  $X$  is an  $n \times p$  matrix (including a column of ones for the intercept),  $Z$  is an  $n \times q$  matrix,  $\beta$  is a  $p \times 1$  vector, and  $\gamma$  is a  $q \times 1$  vector. I assume further that the usual assumptions of the linear model hold conditionally on  $X$  and  $Z$ . That is,  $E(\nu|X, Z) = 0$  and  $V(\nu|X, Z) = \sigma_v^2 I$  where  $I$  is the  $n \times n$  identity matrix. The aim is to make inferences about  $\delta = \beta^* - \beta$ , where  $\beta^*$  is the  $p \times 1$  vector of population least squares regression coefficients of  $Y$  on  $X$  alone. We estimate  $\delta$  with  $d = b^* - b$ , where  $b^*$  is the sample least squares estimator of  $Y$  on  $X$  alone and  $b$  is the least squares estimator of  $Y$  on  $X$  when  $Z$  is included. Let  $g$  be the least squares estimator of  $\gamma$  when both  $X$  and  $Z$  are included in the model, and let  $H = (X^T X)^{-1} X^T Z$ , the  $p \times q$  matrix of least squares regression coefficients for  $Z$  on  $X$ .

In the appendix I show that an appropriate estimator for  $V(d)$  is

$$\hat{V}(d) = H\hat{V}(g)H^T + g^T Wg(X^T X)^{-1}, \quad (10)$$

where  $W$  is a matrix of variances and covariances of the residuals from regressing  $Z$  on  $X$ . The first component of the sum in (10) is equivalent to CPH's conditional variance estimator. Although the matrix  $W$  is *not* standardly available in regression programs, it could be easily computed from the residuals.

No special computation is necessary when there is only a single  $Z$  variable, as (10) reduces to

$$\hat{V}(d) = s^2(g)hh^T + g^2\hat{V}(h), \quad (11)$$

where  $s^2(g)$  is the squared estimated standard error of  $g$ ,  $h$  is now a  $p \times 1$  vector of regression coefficients for  $Z$  on  $X$ , and  $\hat{V}(h)$  is the usual estimated covariance matrix. As in the three-variable case, one may wish to use a more robust estimator for this matrix. From the main diagonal of (11), the variance for the change in the coefficient of a single variable  $X_i$  is

$$s^2(d_i) = s^2(g)h_i^2 + g^2s^2(h_i), \quad (12)$$

where  $h_i$  is the coefficient of  $X_i$  in the regression of  $Z$  on  $X$ . Note that (12) has exactly the same form as (7) in the three-variable case, except that the bivariate coefficient of  $Z$  on  $X$  is replaced by a partial coefficient.

To test the null hypothesis that  $\delta = 0$ , i.e., that *none* of the coefficients changes with the addition of  $Z$ , one may use the statistic

$$F^* = \frac{d^T \hat{V}(d)^{-1} d}{p}, \quad (13)$$

which has approximately an  $F$  distribution with  $p$  and  $n - p - q$  degrees of freedom under the null hypothesis. By contrast, CPH's  $F$ -statistic has  $\min(p, q)$  as the divisor and the corresponding degrees of freedom. That's because CPH's estimator of  $V(d)$  does not have full rank if  $q < p$ , that is, if the number of  $Z$ 's is less than the number of  $X$ 's.<sup>4</sup> But (10) has full rank (except when  $g = 0$ ) with probability 1, because the second component of the sum has full rank.

Likelihood tests of  $\delta = 0$  are also possible with LISREL 8 (Jöreskog and Sörbom 1993) because that program allows for testing nonlinear restrictions on the coefficients. The trick is to specify a recursive linear model in which all the  $X$ 's are causally prior to all the  $Z$ 's. Then estimate the model under the restriction  $Hg = 0$ . While earlier versions of

<sup>4</sup> A square matrix has full rank if no column (or row) is a linear function of the other columns (or rows).

LISREL cannot impose nonlinear restrictions, they do report standard error estimates for the "indirect effects" using formulas given by Sobel (1982). Under a recursive model, the indirect effect of  $X_i$  on  $Y$  through  $Z$  is equivalent to  $d_i$ . Hence, these standard errors can be used to construct  $t$ -statistics and confidence intervals for the individual elements of  $d$ . Although Sobel's formulas are equivalent to (10), his results were obtained as asymptotic approximations. Equation (10), on the other hand, is based on exact results in the appendix.

#### WHEN IS CONDITIONING JUSTIFIED?

I have argued that when predictors are random, making inferences conditional on the sample values of the predictors is justified in some situations but not in others. Is there any way to characterize the distinction between these two kinds of situations? Here are two possible answers to this question:

1. *Ancillarity*.—For conventional inferences about regression coefficients, the practice of making inferences conditional on the predictor variables is well grounded in the theory of ancillarity (Cox and Hinkley 1974). That theory says, roughly, that if the marginal probability distribution of some sample statistic does not depend on the parameters of interest, then inference should be made conditional on that sample statistic, which is termed an ancillary statistic. In standard regression problems, the moments (means, variances, and covariances) of the predictor variables are ancillary for testing hypotheses about the regression coefficients, so it is reasonable to make inferences conditional on those moments.

For the three-variable case considered here, however, we have a sample statistic  $s_{xx}$  whose sampling distribution obviously depends on the parameter  $\sigma_{xx}$ . But this is one of the parameters we are trying to make inferences about because the null hypothesis is true if  $\sigma_{xx} = 0$ . Conditioning on  $s_{xx}$  is tantamount to assuming that the only way the null hypothesis (that  $\delta = 0$ ) can be true is if  $\beta_{yx \cdot z} = 0$ . Hence, conditioning cannot be justified by the principle of ancillarity.

2. *Conditional unbiasedness*.—Consider an arbitrary statistic  $\hat{\theta}$  computed from data  $(X, Y)$ . A plausible claim is that the conditional variance  $V(\hat{\theta}|X)$  is acceptable for making inference about  $\hat{\theta}$  whenever  $V(\hat{\theta}) = E(V(\hat{\theta}|X))$ , that is, when the unconditional variance is equal to the expected value of the conditional variance. From (A3) in the appendix, we see that this equality holds if and only if  $E(\hat{\theta}|X)$  does not depend on  $X$ . For a standard, three-variable linear model, this condition is satisfied because  $E[b_{yx \cdot z}|X, Z] = \beta_{yx \cdot z}$ , that is, the coefficients are conditionally unbiased. On the other hand,

$$E[d|X, Z] = \delta \left( \frac{b_{xx}}{\beta_{xx}} \right) \neq \delta. \quad (14)$$

This means that, in a particular sample, if  $b_{xx}$  happens to be larger (or smaller) than the "true" coefficient due to sampling variation, then  $d$  will also tend to be too large (or too small). Using the unconditional variance for inference enables us to take account of this additional source of variation.

#### GENERALIZED LINEAR MODELS

Everything I have said so far applies only to linear models. For generalized linear models, the situation is rather more complicated because the regression estimators cannot usually be expressed in closed form. Nevertheless, given the arguments just presented, there are strong reasons to suspect that the unconditional variance of  $d$  will exceed the conditional variance. While I have not yet attempted to derive the unconditional variance, I am not optimistic that useful formulas will be forthcoming. Bootstrap estimates of the variance offer one potential solution in these situations.

#### CONCLUSION

I have argued that the methods of CPH for comparing regression coefficients in "full" and "reduced" models depend on treating the predictor variables as though they were fixed from sample to sample, an unrealistic assumption for most social science data. Allowing for randomness in the predictors leads to different, more conservative tests and confidence intervals. This differs from inferences about coefficients in a single, correctly specified model, where methods are the same whether predictors are fixed or random. The alternative methods proposed here are readily implemented with conventional software and appear to outperform the CPH methods on simulated data. While the issues are admittedly subtle, I hope that these arguments are sufficient to forestall the universal and uncritical adoption of the CPH tests as a standard for evaluating changes in regression coefficients.

#### APPENDIX

##### Derivation of the Unconditional Variance

Using the notation developed above for the multivariable case, I make use of the well-known result (e.g., Goldberger 1991) that  $b^* = b + Hg$ . This implies that  $d = Hg$ . We then have



$$E(d|X, Z) = H\gamma, \quad (\text{A1})$$

and

$$V(d|X, Z) = HV(g)H^T. \quad (\text{A2})$$

Formula (A2) is equivalent to the variance formula given by CPH. To get the unconditional variance, we make use of another well-known formula for arbitrary random vectors  $U$  and  $V$ ,

$$V(U) = E[V(U|V)] + V[E(U|V)], \quad (\text{A3})$$

that is, the unconditional variance is equal to the expectation of the conditional variance plus the variance of the conditional expectation. Applying this to the problem at hand, we have

$$\begin{aligned} V(d) &= E[V(d|X, Z)] + V[E(d|X, Z)] \\ &= E(HV(g)H^T) + V(H\gamma). \end{aligned} \quad (\text{A4})$$

Focusing on the second term in (A4), write  $H = [h_1 h_2 \dots h_q]$ ,  $\gamma = [\gamma_1 \gamma_2 \dots \gamma_q]^T$  and  $Z = [Z_1 Z_2 \dots Z_q]$ . We then have

$$H\gamma = \sum_{j=1}^q \gamma_j h_j. \quad (\text{A5})$$

It follows that

$$V(H\gamma) = \sum_j \gamma_j^2 V(h_j) + \sum_j \sum_{k \neq j} \gamma_j \gamma_k \text{cov}(h_j, h_k). \quad (\text{A6})$$

Substituting (A6) into (A4) yields the desired result.

If we are willing to assume that the population regression of  $Z$  on  $X$  satisfies a standard linear model, then we can further simplify (A6). In that event,

$$V(h_j) = \sigma_j^2 E(X^T X)^{-1}, \quad (\text{A7})$$

and

$$\text{cov}(h_j, h_k) = \sigma_{jk} E(X^T X)^{-1}, \quad (\text{A8})$$

where  $\sigma_j^2$  is the disturbance variance in the equation for  $Z_j$  and  $\sigma_{jk}$  is the covariance between the disturbances in the equations for  $Z_j$  and  $Z_k$ . Let  $\Omega$  be a  $q \times q$  matrix containing these variances and covariances. We can then write  $V(d)$  as

$$V(d) = E(HV(g)H^T) + \gamma^T \Omega \gamma E(X^T X)^{-1}. \quad (\text{A9})$$

To estimate this, we can replace population quantities with sample estimators to get

$$\hat{V}(d) = H\hat{V}(g)H^T + g^T Wg(X^T X)^{-1}, \quad (\text{A10})$$

where  $W$  is a matrix of variances and covariances of the residuals obtained from regressing  $Z$  on  $X$ .

## REFERENCES

- Binkley, James K., and Philip C. Abbot. 1987. "The Fixed X Assumption in Econometrics: Can the Textbooks be Trusted?" *American Statistician* 41:206–14.
- Breiman, Leo, and Philip Spector. 1992. "Submodel Selection and Evaluation in Regression: The X-Random Case." *International Statistical Review* 60:291–319.
- Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. 1995. "Statistical Methods for Comparing Regression Coefficients between Models." *American Journal of Sociology*. In this issue.
- Cox, D. R., and D. V. Hinkley. 1974. *Theoretical Statistics*. Cambridge: Chapman & Hall.
- Goldberger, Arthur S. 1991. *A Course in Econometrics*. Cambridge, Mass.: Harvard University Press.
- Joreskog, Karl G., and Dag Sorbom. 1993. *New Features in LISREL 8*. Chicago: Scientific Software.
- Sobel, Michael E. 1982. "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models." Pp. 290–312 in *Sociological Methodology 1982*, edited by Samuel Leinhardt. San Francisco: Jossey-Bass.
- White, Halbert. 1980. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica* 48:817–38.

## REPLY TO ALLISON: MORE ON COMPARING REGRESSION COEFFICIENTS<sup>1</sup>

Allison (1994) raises some legitimate questions concerning the interpretation of some of our procedures (Clogg, Petkova, and Haritou 1994; CPH hereafter). Because his comments pertain to the CPH methods for the analysis of linear regression models, in this note we also concentrate on this case.<sup>2</sup>

Virtually all of the literature on specification tests or collapsibility tests

<sup>1</sup> We are indebted to Bing Li and Bruce Lindsay for helpful comments.

<sup>2</sup> The CPH procedures for generalized linear models are different from the CPH procedures for linear models. For example, for the analysis of collapsibility in contingency tables where log-linear models or related models might be used, the CPH procedures are automatically "unconditional" because the models pertain to multivariate (joint or marginal) distributions. For contingency table settings where either log-linear models (unconditional) or logit models (conditional) might be used to address the same question, the inferences are equivalent between the two cases as well. The CPH procedures for generalized linear models are valid unconditionally, for testing the relevant null hypothesis, and so are valid conditionally as well.

cited in CPH pertains to unconditional inference and large-sample approximations. Because the CPH methods are related to these other methods, the CPH methods also have valid unconditional interpretations. In this note we describe the situations under which the CPH methods (for the analysis of linear models) have valid unconditional interpretations. We also give a simple modification of those methods, based directly on formulas in CPH, that can be used to obtain unconditional tests of the null hypotheses in all possible situations.

Our summary responses to the questions in Allison's comment are as follows:

1. For fixed predictors, the CPH procedures are valid and optimal in a variety of senses, under standard assumptions. In addition to CPH, see Clogg, Petkova, and Shihadeh (1992).

2. For random predictors, the procedures in CPH are valid when inferences are "conditional" on the relevant functions of the predictor values observed. The CPH test can be interpreted as giving the significance (or the approximate significance) of observed between-model differences in regression coefficients, assessed in terms of repeated sampling, with repetitions having the same (or similar) values of the predictors observed in the sample (i.e., essentially as hypothetical repeated sampling of "error terms").

3. For random predictors, the CPH procedures give valid unconditional inferences for testing the null hypothesis in some but not all cases. The CPH procedures do not provide valid (unconditional) inferences when the two sets of predictors are orthogonal (or nearly orthogonal) in the population. A simple modification, based directly on the formulas in CPH and described below, gives valid unconditional inference in all cases that give rise to the null hypothesis.

## I. THE PROBLEM

The (linear-model) problem considered in CPH can be described in matrix notation as follows. Consider a "reduced" model

$$H_R: Y = X\beta^* + \epsilon, \quad (1)$$

and a "full" model

$$H_F: Y = X\beta + Z\gamma + \nu. \quad (2)$$

Suppose that all variables are measured as deviations from means so that intercepts can be ignored. When viewed unconditionally, it is assumed that  $\nu$  is uncorrelated with the predictors  $X$  and  $Z$ , which is one way of stating that model  $H_F$  is assumed to be the "true" model. As in CPH

we seek inference procedures that are logically consistent with this assumption.

The *null hypothesis* of interest is

$$H_0: \beta^* = \beta. \quad (3)$$

Using (1) and (2) we obtain  $\epsilon = Z\gamma + X(\beta - \beta^*) + v$ , and when the null hypothesis is true we obtain

$$\epsilon = Z\gamma + v. \quad (4)$$

Using (4) we obtain another representation of the null,

$$\Sigma_{XZ}\gamma = 0, \quad (5)$$

where  $\Sigma_{XZ}$  is the matrix of covariances between predictors in  $X$  and predictors in  $Z$ .<sup>3</sup> In other words, the null hypothesis can be viewed as the condition  $\beta^* = \beta$ , or as the condition that the error term  $\epsilon$  is uncorrelated with  $X$  or as the condition in (5).<sup>4</sup> From (5) we see that  $H_0$  is true if any one of the following conditions holds:

$$\Sigma_{XZ} = 0, \gamma \neq 0, \quad (i)$$

( $X$  and  $Z$  are orthogonal in the population) or

$$\gamma = 0, \Sigma_{XZ} \text{ unspecified}, \quad (ii)$$

( $Z$  does not predict net of  $X$ ) or

$$\Sigma_{XZ} \neq 0, \gamma \neq 0, \Sigma_{XZ}\gamma = 0, \quad (iii)$$

(rows of  $\Sigma_{XZ}$  are orthogonal to  $\gamma$ ).<sup>5</sup>

## II. ESTIMATING VARIANCES

Let  $b^*$  denote the least squares estimator of  $\beta^*$  in the reduced model, and let  $b$  denote the least squares estimator of  $\beta$  in the full model. The variance of  $d = b^* - b$  can be written as

$$V(d) = V(b^*) + V(b) - 2 \text{cov}(b^*, b). \quad (6)$$

This variance can be understood as either conditional or unconditional.

<sup>3</sup> The "population" regression coefficient  $\beta^*$  can be written as  $\beta^* = \Sigma_{XX}^{-1}\Sigma_{XY} = \beta + \Sigma_{XX}^{-1}\Sigma_{XZ}\gamma$ , assuming that  $v$  and  $X$  are uncorrelated, and hence  $\beta^* = \beta$  if and only if  $\Sigma_{XZ}\gamma = 0$ . There are several other ways to represent the null hypothesis.

<sup>4</sup> It is easy to see that  $V(\epsilon) = \sigma_v^2 = \gamma^T \Sigma_{ZZ} \gamma + (\beta - \beta^*)^T \Sigma_{XX} (\beta - \beta^*) + \sigma_v^2 + 2(\beta - \beta^*)^T \Sigma_{XZ} \gamma$ . When  $H_0$  is true, this simplifies to  $\sigma_v^2 = \gamma^T \Sigma_{ZZ} \gamma + \sigma_v^2$ .

<sup>5</sup> These three cases are mutually exclusive. Case (ii) is the most inclusive, while case (iii) is the least inclusive. Note that case (iii) is not relevant when there is just one  $X$  variable and just one  $Z$  variable, as in Secs. II and III of CPH.

Use subscript  $u$  for unconditional and subscript  $c$  for conditional, where conditional throughout means "conditional on  $X$  and  $Z$ ." For standard linear models, CPH shows that  $\text{cov}_c(b^*, b) = V_c(b^*)$ , so from (6) we obtain

$$V_c(d) = V_c(b) - V_c(b^*), \quad (7)$$

which is derived in the appendix of CPH. For conditional inference (i.e., conditional on  $X$  and  $Z$ ), this variance should be evaluated using the mean-squared error from the full model, which is another result in CPH. The sample estimators are

$$\hat{V}_c(b^*) = \hat{\text{cov}}_c(b^*, b) = (X^T X/n)^{-1} \hat{\sigma}_v^2/n; \quad (8)$$

$$\hat{V}_c(b) = (nA) \hat{\sigma}_v^2/n, \quad (9)$$

where  $nA = [(X^T X/n) - (X^T Z/n)(Z^T Z/n)^{-1}(Z^T X/n)]^{-1}$  (see eq. [A6] in CPH) and  $\hat{\sigma}_v^2$  is the mean-squared error from the full model.

Note that if  $X$  and  $Z$  are orthogonal in the population, in the limit  $(X^T Z/n)$  approaches  $\Sigma_{XZ} = 0$ , and using (8), (9), and the definition of the matrix  $A$ , we see that  $\hat{V}_c(b^*)$  and  $\hat{V}_c(b)$  tend to the same value. In other words,  $\hat{V}_c(d) = \hat{V}_c(b) - \hat{V}_c(b^*)$  approaches zero quickly under this condition, so the conditional variance estimator  $\hat{V}_c(d)$  is simply not appropriate if this condition obtains.<sup>6</sup> In view of these facts, the situation where the null is true because  $X$  and  $Z$  are orthogonal and  $\gamma \neq 0$  (case [i]) must be excluded in the conditional approach used in CPH, if these procedures are to be used for unconditional inference.<sup>7</sup>

Now consider the unconditional variance of  $d$ , that is,  $V_u(d) = V_u(b^*) + V_u(b) - 2 \text{cov}_u(b^*, b)$ . The unconditional covariance in question is consistently estimated by  $\hat{\text{cov}}_c(b^*, b) = (X^T X/n)^{-1} \hat{\sigma}_v^2/n$ , because the covariance is an expectation, and  $(X^T X/n)^{-1}$  tends to  $\Sigma_{XX}^{-1}$ , which is

<sup>6</sup> Because  $(X^T Z/n)$  and its transpose both appear in a product of terms in the expression for  $\hat{V}_c(b)$ ,  $\hat{V}_c(d)$  tends to zero rapidly when  $X$  and  $Z$  are orthogonal in the population; i.e., it tends to zero more rapidly than  $\sigma_v^2/n$  approaches zero. The rate of convergence to zero is essentially  $1/n^2$  when this condition is satisfied. This explains Allison's results in the first row of his table of simulations. When  $\Sigma_{XZ} \neq 0$ , the quantity converges to zero at the usual rate,  $1/n$ . The asymptotic variance estimator in question converges to  $[(\Sigma_{XX} - \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{ZX})^{-1} - \Sigma_{XX}^{-1}]\sigma_v^2/n$ , and if  $X$  and  $Z$  are orthogonal then this quantity is identically zero. If  $\Sigma_{XZ} = 0$  and  $\gamma = 0$ , as in case [ii], then this conditional variance estimator is valid unconditionally, however.

<sup>7</sup> Note that an analogous restriction applies if the CPH methods are viewed as conditional procedures (see nn. 2 and 4, CPH). The related Hausman (1978) test, which is an unconditional procedure, is also valid only if  $X$  and  $Z$  are not orthogonal in the population. The Hausman estimator, if adapted for use in this situation, merely replaces  $\sigma_v^2$  by  $\sigma_u^2$ , where  $\sigma_u^2$  is the error variance under the reduced model. The corresponding asymptotic variance (Hausman method) will be identically zero when this type of orthogonality holds also.

bounded under standard assumptions (variances are finite, predictors in  $X$  are not colinear). Similarly, the unconditional variance of  $b$  is consistently estimated by  $\hat{V}_c(b) = (nA)\hat{\sigma}_v^2/n$ , which is the usual estimator of the variance of  $b$  in the full model. The unconditional variance of  $b^*$  is not always consistently estimated by  $\hat{V}_c(b^*)$ , however. But  $V_*(b^*)$  can be consistently estimated by

$$\hat{V}_*(b^*) = (X^T X/n)^{-1} \hat{\sigma}_e^2/n, \quad (10)$$

where  $\hat{\sigma}_e^2$  is the mean-squared error from *the reduced model*. Note that  $\hat{V}_*(b^*)$  is the usual estimator of the variance of  $b^*$ , as obtained from estimating the reduced model. When the null is true,  $b^*$  estimates  $\beta$ , but when the null is not true,  $b^*$  is biased (or inconsistent) as an estimator of  $\beta$ ; in this case  $b^*$  estimates  $\beta^* (\neq \beta)$ . When the null is not true,  $X$  and  $\epsilon$  will be correlated, but  $\hat{V}_*(b^*)$  is still a consistent estimator of the variance of  $b^*$ .<sup>8</sup>

### III. A SIMPLE UNCONDITIONAL VARIANCE ESTIMATOR AND A TEST

The above results imply the following estimator for the (unconditional) variance of  $d$ . Using (8), (9), and (10), we obtain

$$\hat{V}_*(d) = \hat{V}_*(b^*) + \hat{V}_c(b) - 2\hat{V}_c(b^*). \quad (11)$$

Note that  $\hat{V}_c(b)$  and  $\hat{V}_c(b^*)$  are consistent estimators of  $V_*(b)$  and  $\text{cov}_*(b^*, b)$ , respectively, and that  $\hat{V}_*(b^*)$  defined in (10) is a consistent estimator of  $V_*(b^*)$ . So  $\hat{V}_*(d)$  is a consistent estimator of the (unconditional) variance of  $d$ .<sup>9</sup>

For the  $k$ th predictor  $X_k$ , the squared standard error of the difference  $d_k = b_k^* - b_k$  can be estimated by

$$s_u^2(d_k) = s^2(b_k^*) + s^2(b_k) - 2s^2(b_k^*)(\hat{\sigma}_v^2/\hat{\sigma}_e^2) \quad (12)$$

using (11). Only the (squared) standard errors and the mean-squared errors from the two models are used:  $s^2(b_k^*)$  is the squared standard error of  $b_k^*$  (as calculated in the reduced model), and  $s^2(b_k)$  is the squared standard error of  $b_k$  (as calculated in the full model). The significance of  $d_k$  can be assessed, in large samples, by  $z = d_k/s_u(d_k)$ , which follows a

<sup>8</sup> The fact that  $X$  and the error  $\epsilon$  will be correlated when the null is not true does not create a problem. It is still the case that  $E(b^*|X) = \beta^*$ ,  $V(b^*|X) = (X^T X)^{-1} \sigma_e^2$ . And  $V(b^*) = E_X[V(b^*|X)]$ —because  $V_X[E(b^*|X)] = V_X(\beta^*) = 0$ . In the usual errors-in-variables problem, when the observed predictors are correlated with the error term, these relationships still hold, but of course  $E(b^*|X)$  can be very different from  $E(b^*|X^*)$ , where  $X^*$  is the “true” value of the predictor.

<sup>9</sup> The CPH method for generalized linear models is analogous to the expression in (11) above; see CPH eqq. (22)–(25).

standard unit normal under the null.<sup>10</sup> The variance estimator in (11) or (12) and the (large-sample) test based on it are valid for all cases that give rise to the null hypothesis (cases [i], [ii], or [iii]).

The CPH procedures for conditional inference, in contrast, give

$$s_c^2(d_k) = s^2(b_k) - s^2(b_k^*)(\hat{\sigma}_v^2/\hat{\sigma}_e^2) \quad (13)$$

(see CPH, eq. [18]). Note that  $s_{\mathbf{u}}^2(d_k) \geq s_c^2(d_k)$ , in large samples. Using (12) and (13), we find that  $s_{\mathbf{u}}^2(d_k) = s_c^2(d_k) + s^2(b_k^*)[1 - \hat{\sigma}_v^2/\hat{\sigma}_e^2]$ , and when  $\gamma = 0$  the two mean-squared errors will be approximately equal so that the extra component can be ignored. Using the above definitions of variances and covariances, it is easy to verify that the unconditional correlation between  $b_k^*$  and  $b_k$  is related to the conditional correlation as  $\hat{\text{corr}}_{\mathbf{u}}(b_k^*, b_k) = [\hat{\text{corr}}_c(b^*, b)](\hat{\sigma}_v/\hat{\sigma}_e)$ . Because  $\sigma_v \leq \sigma_e$ , in large samples the unconditional correlation in question will tend to be smaller in absolute value than the conditional correlation, but the two correlations will be equal, in large samples, if  $\gamma = 0$ .

#### IV. UNCONDITIONAL INTERPRETATIONS OF THE CPH CONDITIONAL PROCEDURE

As demonstrated above, the "problem" with the CPH conditional procedure is that, in some cases, it can lead to inconsistent estimation of the unconditional variance of  $b^*$ , the least squares estimator of the regression coefficient in the reduced model.

The unconditional variance of  $b^*$  can be written as

$$V_{\mathbf{u}}(b^*) = E_{X,Z}[V_c(b^*)] + V_{X,Z}[E(b^*|X, Z)], \quad (14)$$

which upon direct substitution gives

$$V_{\mathbf{u}}(b^*) = E_{X,Z}[(X^T X/n)^{-1} \sigma_v^2/n] + V_{X,Z}[(X^T X/n)^{-1} (X^T Z/n) \gamma]. \quad (15)$$

Note that the first component of  $V_{\mathbf{u}}(b^*)$  is consistently estimated by  $\hat{V}_c(b^*)$ , as defined in (8). And note further that  $V_{\mathbf{u}}(b^*)$  is consistently estimated by  $\hat{V}_{\mathbf{u}}(b^*)$  defined by (10). Straightforward algebra using (6), (7), and (14) gives

$$V_{\mathbf{u}}(d) = E_{X,Z}[V_c(d)] + V_{X,Z}[(X^T X/n)^{-1} (X^T Z/n) \gamma],$$

which is derived in a different way by Allison. Note that the term  $V_{X,Z}(\cdot)$  can be ignored when the null hypothesis is true because  $\gamma = 0$  (case [ii]), in which case this term is identically zero. This term will be negligible (and hence can be ignored) in some other cases. But this term cannot be

<sup>10</sup> The validity of (12) and the test based on it can be verified through simulations. Each term in (12) is a consistent estimator of the corresponding unconditional variance or covariance, and each term tends to zero at the usual rate  $1/n$ .

ignored if the null hypothesis is true because  $X$  and  $Z$  are orthogonal in the population and  $\gamma \neq 0$  (case [i]).

The CPH procedures thus determine the (unconditional) significance of  $d$  in terms of the significance of  $\hat{\gamma}$ . That is, they are sensitive to condition [ii], which gives rise to the null; but they are not sensitive to condition [i].<sup>11</sup> Allison's simulations under this condition in the bottom half of his table verify this in a simple case. (The CPH conditional procedures are "exact" unconditionally in this case.) The CPH procedures are not valid if the null is true because  $X$  and  $Z$  are orthogonal and  $\gamma \neq 0$ . If this condition is a likely possibility then the simple modification presented above can be used.

## V. CONCLUSION

There can be a difference between conditioning and not conditioning when the predictors are random. Whether or not to condition on sample values of random predictors depends on the manner in which the data were collected and on what inferences are sought. When unconditional inference is appropriate, it is important to take account of the ways in which the null hypothesis could arise.

Researchers should use the CPH conditional procedures if conditional inferences are appropriate. Experimental data and many types of quasi-experimental data are suited for the conditional procedure. For the regression analysis of naturally occurring observations (say, observations on states or organizations), it seems most natural to "fix" the values of the predictor variables. Inferences would then be based on (hypothetical) repetitions of error terms.

Researchers should use the CPH conditional procedure with an unconditional interpretation if unconditional procedures are appropriate (and orthogonality between  $X$  and  $Z$  can be ruled out). In this situation researchers should recognize that the significance of the between-model difference in regression coefficients is then judged on the basis of the significance of  $\hat{\gamma}$ , or that the inferences are sensitive to case (ii) that gives rise to the null. In this case, between-model comparisons of regression coefficients must be construed primarily as a function of the significance of the added ( $Z$ ) variables. We think this situation might characterize many practical problems in social research where two explanations are compared. The need to compare explanations might arise most often because the two sets of predictors involved are "known" to be correlated.

<sup>11</sup> Suppose  $X$  has  $p$  predictors and  $Z$  has  $q$  predictors. If  $p \geq q$ , the CPH test for the significance of all  $p$  elements of  $d$  is equivalent to the usual  $F$ -test for the significance of the incremental contribution of  $Z$ . That is, the CPH test and the usual  $F$ -test for the incremental contribution of  $Z$  are closely related, in the sense that the former is always a function of the latter.



Finally, researchers should use the unconditional procedure given in Section III of this note if unconditional procedures are appropriate and all possible ways that the null can be true are of interest. The CPH procedures, including the simple modification presented in this note (see eqq. [11] and [12]), are very simple to apply, and we believe that these procedures solve an important problem.

CLIFFORD C. CLOGG  
*Pennsylvania State University*

EVA PETKOVA  
*Columbia University*

with the assistance of  
TZUWEI CHENG  
*Pennsylvania State University*

#### REFERENCES

- Allison, Paul D. 1994. "The Impact of Random Predictors on Comparisons of Coefficients between Models." *American Journal of Sociology*. In this issue.
- Clogg, Clifford C., Eva Petkova, and Edward S. Shihadeh 1992. "Statistical Methods for Analyzing Collapsibility in Regression Models." *Journal of Educational Statistics* 17:51-74.
- Hausman, Jerry A. 1978. "Specification Tests in Econometrics." *Econometrica* 50: 1251-71.

## Commentary and Debate

To conserve space for the publication of original contributions to scholarship, the comments in this section must be limited to brief critiques. They are expected to address specific errors or flaws in articles and reviews published in the *AJS*. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Longer or less narrowly focused critiques should be submitted as articles. Authors of articles and reviews are invited to reply to comments, keeping their replies to the length of the specific comment. The *AJS* does not publish commenters' rebuttals to authors' replies. We reserve the right to reject inappropriate or excessively minor comments.

### THE STUDY OF SOCIAL TASTE THROUGH FIRST NAMES: COMMENT ON LIEBERSON AND BELL

The article "Children's First Names: An Empirical Study of Social Taste" (*AJS* 98 [November 1992]: 511–54) by Stanley Lieberman and Eleanor O. Bell raises two problems: (1) it ignores previously published French-language studies dealing with exactly the same topic and (2) the findings presented in the article are limited. These two problems are linked, for had Lieberman and Bell presented the findings of previous studies on the same topic, the weakness of their own findings would have been obvious.

It was my understanding that Lieberman and Bell knew those previous works. Why did they not mention them? In their article, Lieberman and Bell deal at length with the idea that first names offer research material particularly well suited to the study of social taste and its transformations. I first presented this idea, nearly in the same terms, in an article published 14 years ago (Besnard 1979). This article is not mentioned, although I assume that Lieberman knew it, for I sent it to him at his request in 1988.

This article outlined a program and contributed to the creation of a research tradition on first names in France to which historians, demographers, and sociologists have contributed. This tradition of research is

Permission to reprint a comment printed in this section may be obtained only from the author.

illustrated by several articles (mentioned below) and no less than four books (Dupâquier et al. 1984; Pérouas et al. 1984; Besnard and Desplanques 1986; Dupâquier, Pélissier, and Rébaudo 1987). Although Lieberman and Bell mention the name of one of these books (Besnard and Desplanques 1986), nothing is said about its content. Of the broader research tradition, only passing mention is made: "In addition, there are some sociological and demographic studies of naming outside the United States" (p. 513n.3). Is this localism justified?

The problem of excluding these French studies is that, in fact, they have been carried out on a much larger scale than Lieberman and Bell's own work. Pérouas et al. (1984), for example, analyze first names in a French province for a period of 1,000 years ( $N = 1,000,000$ ). Dupâquier et al. (1987) analyze first names throughout France during the entire 19th century from a representative sample of 90,000 persons. A third study (Besnard and Desplanques 1986) uses several national samples representing a total of 2.5 million individuals (4 million in the fourth edition of the same book) to analyze first names given in France from 1890 to 1985. In the latter study, a specifically sociological approach is used to correlate the first names of all the individuals born since 1940 with a wide range of social and demographic variables such as birth order, parents' age, and father's, mother's, and grandparents' occupation. These studies, which are well known in France, present findings concerning social tastes and the social transformation of tastes whose validity has been widely accepted.

In addition to the books already mentioned, these findings have continued to be elaborated and discussed in articles (Besnard 1984; Desplanques 1986*a*; Bozon 1987*a*; Besnard 1991; Besnard and Grange 1993; Besnard 1994) as well as in pedagogical instruments such as encyclopedias and dictionaries (Boudon et al. 1989; Besnard 1988; Besnard 1989) and handbooks (Cherkaoui 1992). Results of the French naming studies have also been discussed in the main French journals of social sciences (Besnard 1986; Desplanques 1986*b*; Bozon 1987*b*; Héran 1987; Mendras 1987; Pérouas 1987) and debated in symposia (Dupâquier et al. 1984).

Why did Lieberman and Bell not take into account these studies that focused on the same sociological issue and used the same kind of material? It is difficult to imagine that the authors wanted to enhance the originality of their work.

The question of priority is not all that important. It is quite possible that the idea of using first names as a means for exploring social taste may have occurred to Lieberman and Bell spontaneously, albeit recently, for it did not appear in Lieberman's initial programmatic article on first names (Lieberman 1984).

More important is the fact that, compared with French studies, the

findings presented by Lieberman and Bell are rather limited. The authors themselves admit these limitations, but, according to them, they are due to the pioneering nature of their work which, supposedly, deals with "a relatively neglected topic" (p. 549).

As a matter of fact, most of their findings are not established on a solid enough basis. The reason is simple: the data sets they use are too limited, and, in particular the time span taken into account (13 years) is far too brief. In such a short period, it is impossible to grasp the life cycle of a first name and to know which moment in this cycle we are observing: Is the first name just appearing, is it developing, at its apex, declining, disappearing? This point has an important consequence. We learn from Lieberman and Bell that, in New York State during the period 1973–85, certain first names are more likely to be chosen by mothers with a high school education than by mothers with only an elementary school education. But from this fact alone one cannot draw the conclusion, as they have done, that certain first names are typical of certain social classes. Indeed, the differences registered could be explained not so much by "class differences in taste" as by the existence of a vertical pattern of diffusion of tastes that entails a temporal gap in the choice of first names according to the social status of the parents. I refer here to the famous trickle-down process, which have been put forward, for more than a century, by sociologists, but rarely studied thoroughly for itself.

In this respect, Lieberman and Bell remain silent on essential information which I had to check myself.<sup>1</sup> They described a number of first names—such as Emily, Lauren, Megan, Andrew, Matthew, Adam, Ryan, Joshua—as the most favored by the highly educated, but they forget to mention that these same first names were in an ascending phase during the period 1973–85 in New York State. Conversely, first names such as Michelle, Melissa, Lisa, Anthony, Robert, Richard, and Jason, described by Lieberman and Bell as typical of the less educated, were in decline during the same period. This means that vertical diffusion process might be a better explanation than class differences.

Strangely enough, the authors do not perceive the contradiction between their strong assertions on class differences in taste and their subsequent attempt to stress the existence of a pattern of vertical diffusion. They admit, on this point, that the "span of time . . . is too brief"; however, they present some findings with the same justification: "To our knowledge, this diffusion issue is rarely studied with actual temporal data" (p. 542). In fact this issue of the social diffusion and transformation of tastes has been at the very center of the French studies.

<sup>1</sup> Each year the New York State Health Department provides the listing of the top 25 first names of both sexes in order of frequency. This information is not very precise but is sufficient to determine whether the use of a first name is increasing or declining.

In order to assert with confidence that a first name has been typical of a certain social group, the use of this first name must be analyzed during its entire career. Whereas, in France, the life span of a first name is 40 years or so, names appear to persist longer in the United States. It is possible, indeed even likely, that while first names such as Lauren or Megan were overrepresented in the upper middle classes in New York State prior to 1985, they became typical of the lower classes after 1985. This possibility undermines Lieberman and Bell's arguments about the "favorable disposition" of the highly educated "toward the *n*-ending name" (p. 531), and, more generally, about "*the nature* of educational differences in taste" (p. 529; my emphasis). Despite their claim to the contrary then, Lieberman and Bell have failed to demonstrate that there are "taste differences in subpopulations" that would reveal their "general esthetics dispositions" (p. 511). I believe this was the main objective of their study.

However, such a demonstration has been made. The French studies, not discussed in Lieberman and Bell's article, have demonstrated both (1) the important influence of the vertical diffusion of tastes and (2) the existence of class differences in taste. Some first names are indeed more or less typical of certain social classes, but many others are not. One can find names of this latter type in all social classes with practically the same frequency, but at different periods.

In order to demonstrate social stratification of taste, it is vital to master the diffusion process and thus cover a long period. This key issue has not been perceived by Lieberman and Bell, although it was abundantly documented in the French studies at their disposal. Have Lieberman and Bell considered that sociological research is restricted to works published in the United States and using U.S. materials? If such was the case, it would be a striking example of scientific parochialism.

PHILIPPE BESNARD

*Université de Paris-Sorbonne*

## REFERENCES

- Besnard, Philippe. 1979. "Pour une étude empirique du phénomène de mode dans la consommation des biens symboliques: Le cas des prénoms." *European Journal of Sociology* 20:343-51.
- . 1984. "De la sous-exploitation des prénoms dans la recherche sociologique." Pp. 51-59 in *Le prénom: Mode et histoire*, edited by J. Dupâquier, A. Bideau, and M.-E. Ducreux. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.
- . 1986. Review of Pérouas et al., 1984. *Revue française de Sociologie* 27:340-42.

- . 1988. "Le choix des prénoms." Pp. 383–87 in *Universalis 1988*. Paris: *Encyclopaedia Universalis*.
- . 1989. "La mode dévêtuée: Travaux empiriques." *Encyclopaedia Universalis* 15:528–29.
- . 1991. "Le choix d'un prénom: Actualité de la méthode Durkheimienne." *Recherches Sociologiques* 22:53–60.
- . 1994. "A Durkheimian Approach to the Study of Fashion: The Sociology of First Names." Pp. 159–73 in *Debating Durkheim*, edited by W. S. F. Pickering and H. Martins. London: Routledge.
- Besnard, Philippe, and Guy Desplanques. 1986. *Un prénom pour toujours: La cote des prénoms*. Paris: Balland.
- Besnard, Philippe, and Cyril Grange. 1993. "La fin de la diffusion verticale des goûts? Prénoms de l'élite et du vulgum." *L'Année Sociologique* 43:269–94.
- Boudon, Raymond. 1989. *Dictionnaire de la sociologie*. Paris: Larousse.
- Bozon, Michel. 1987a. "Histoire et sociologie d'un bien symbolique, le prénom." *Population* 42:83–98.
- . 1987b. "Deux ouvrages sur les prénoms." *Population* 42:561–63.
- Cherkaoui, Mohamed. 1992. "Stratification." Pp. 97–152 in *Traité de sociologie*, edited by R. Boudon. Paris: Presses Universitaires de France.
- Desplanques, Guy. 1986a. "Les enfants de Michel et Martine Dupont s'appellent Nicolas et Céline." *Economie et Statistique* 184:63–83.
- . 1986b. Review of Dupâquier et al. 1984. *Revue Française de Sociologie* 27:339–40.
- Dupâquier, Jacques, Alain Bideau, and Marie-Elizabeth Ducreux, eds. 1984. *Le prénom: Mode et histoire*. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.
- Dupâquier, Jacques, Jean-Pierre Pélissier, and Danièle Rébaudo. 1987. *Le temps des Jules: Les prénoms en France au XIXe siècle*. Paris: Christian.
- Héran, François. 1987. Review of Besnard and Desplanques, 1986. *Revue Française de Sociologie* 28:708–11.
- Lieberson, Stanley. 1984. "What's in a Name? . . . Some Sociolinguistic Possibilities." *International Journal of the Sociology of Language* 45:77–87.
- Mendras, Henri. 1987. Review of Besnard and Desplanques, 1986. *Futuribles* 114:88–91.
- Pérouas, Louis. 1987. Review of Besnard and Desplanques, 1986, and of Dupâquier et al., 1987. *Annales* 42:643–45.
- Pérouas, Louis, et al. 1984. *Léonard, Marie, Jean et les autres: Les prénoms en Limousin depuis un millénaire*. Paris: Éditions du Centre National de la Recherche Scientifique.

## REPLY TO PHILIPPE BESNARD

Besnard raises three objections to our paper: (1) we do not cite the French literature on names; (2) our interpretation of class differences does not emphasize vertical imitation; (3) we do not cite his 1979 paper because we want to claim his ideas as being our own. I will address each of these in the above order.

It is indeed true that there is a long and distinguished history of French scholarship on first names. I have gone through enough of these publications to develop an admiration for the care and elegance involved, as

well as the craftsmanship of the work. Some of these studies should be useful to me as I address cross-national comparisons of changes over time. However, the issue at hand is their relevance to the work by Bell and myself. Why does our paper ignore the French studies, endowed as they are with such huge *N*'s? All I can do is offer a rather prosaic line of reasoning. One does not cite work just to cite it—rather one cites work relevant to the issues under consideration. There are 11 tables in our paper as well as two in the appendix. Table 1 lists the leading boys' and girls' names, separately for blacks and whites. Table 5 compares the influence of race with the influence of parental education on the names given to children, showing that the race effect is stronger than the class effect. Table 11 tabulates the use of names as reported in a list of standard American names with education and race of mother. For each of these three tables, why not cite the relevant French data on racial differences? Alas, I can find no information about them in any of the publications referred to above—not even in Besnard and Desplanques (1993), which reports names for 4 million people. Table 6 provides information on the frequency of made-up names—a phenomenon that is really not comparable in France because there is an official control exerted on name choices. As for the data in table 6 on concentration in the top 20 names, Besnard and Desplanques (1993, pp. 18–21) report only the top 10 in each birth cohort. Tables 7 and 8 apply a modified Boolean approach simultaneously to the linguistic, etymological, biblical, historical, and phonemic character of the names favored by different classes. In turn, table 9 looks at the semantic connotations of names favored by the social classes. This whole line of analysis is rejected out of hand by Besnard, but no data are cited from the French literature as to why it is wrong. Are there no French data on this? If not, how can I cite their relevant studies? If there are French data that show we are wrong, why not specifically cite the relevant analysis? That I would find useful. As for the two tables in the appendix, listing the 50 most popular names among white boys and white girls, I think it would have been fun to compare them with French names for a comparable period. The best I could find during that period in Besnard and Desplanques (1993) is the top 10. Apparently, it is possible to list a larger number of the most common names in the entire population, but, in fairness, this must prevent them from listing the 50 most common names for each specific birth cohort. (I will shortly consider the relevance of the French literature for the other tables in our paper.)

Let us turn to the one substantive issue raised by Besnard—whether the observed class differences in naming preferences are due to imitation as contrasted with differences in taste. Table 2 of our article provides a summary measure of the association between distance in education with distance in naming choices. A very orderly pattern shows up. Likewise,

a very detailed analysis of the educational gradients of names is shown in tables 3 and 4. In the case of Emily, for example, the degree of parental preference varies positively with mother's education. We examine these results in terms of class differences in taste such that some names will differ in their appeal to parents with different educational levels. From this perspective, the challenge is to figure out what it is about those names that causes them to vary in their appeal to parents with different levels of education. Besnard rejects this approach, arguing that these class-linked patterns result from a diffusion process in which lower SES parents imitate the names used by those of higher status; in turn, the latter begin to use newer names in order to avoid being associated with these strata. Hence the association between class and name at a given point in time is nothing more than a function of where each name happens to be along the diffusion chain. If I understand the argument, then names such as Crystal, Tammy, Maria, Angela, Luis, and Jose (which we found to be disproportionately favored by mothers with relatively low education) were actually copied from high SES families that used to favor these names. Keep in mind that these two perspectives are not necessarily exclusionary in the sense that we cannot *a priori* say if one is true the others are false. Class differences among some names may differ because they genuinely differ in their appeal to different classes. Yet, on the other hand, the differential popularity of some names may reflect some variant or another of the trickle-down theory.

In order to evaluate these theories, we must first avoid the confusions running through Besnard's argument. Diffusion and imitation are not the same thing. If there is imitation, then there will be diffusion. But, if there is diffusion, it does not follow that it is caused by imitation. In other words, diffusion is not sufficient evidence that imitation occurs. This is because diffusion can be nonrandom (in the sense that one segment of the population may adopt a name earlier than another segment) without it being caused by imitation. The diffusion will occur because of a central stimulus, say television or movies, that amplifies the choices of writers and agents and thereby influences the choices of others. From this perspective, it is not one class imitating the names that they see used by another class; rather it is being inspired by the names they encounter in the mass media as the names of appealing fictional characters or appealing performers.<sup>1</sup> The mechanisms operating even here are more complex than the simplistic imitation model would suggest. For example, the majority of names in the mass media do not take off as popular names. The imitation notion is in itself unable to account for the successes and

<sup>1</sup> I am indebted to Cleveland Kent Evans for suggesting such a mechanism.



the failures. Moreover, a child may be given the name of some illustrious figure not because they are the member of some class that the parents aspire to, but simply because the parents admire or wish to honor that person. Also keep in mind that the imitation model attaches no significance to the nature of the name initially adopted by the higher status parents—all it has to be is a name that is not associated with the classes beneath them. Likewise, the imitation model attributes a mechanical adoption by lower classes: if classes above them use a name, then that is enough for them to decide to also use the same name.

A second confusion found in Besnard's comment is between the number of years that a name is popular in the society and the number of years it takes for a name adopted by one class to then also be adopted by another class (as posited by the imitation theory). Clearly our span of 13 years is too short to draw conclusions about the life span of most names. Quite properly, we avoid this issue. On the other hand, it turns out that 13 years is not too short a period to deal with the diffusion hypothesis; that is, during this brief period we find names that first become popular for at least one educational class and then get adopted by one or more other classes. The pattern is harmonious with a class diffusion hypothesis in the sense that usage of a name among more highly educated mothers is more likely to precede usage among less educated mothers than are shifts in the opposite direction (see our table 10). Take Jason and Ashley as examples. In each case they were not initially adopted by all classes in the same year, but the span between the first and last adoption is so narrow that it is hard for us to believe a trickle-down model could explain it. For reasons developed in our article (see pp. 543–44), we argue that it is questionable that such an incredibly rapid chain could reflect imitation of one class by another. It is hard to understand how lower SES parents would so rapidly learn about the statistical distribution of names among higher SES parents. (In the case of clothing, it is easy to see how this might occur because of movies, television, magazines, and the clothing shown in stores.) To be sure, there may well be other names, where the diffusion is both slower and in the direction hypothesized, that are consistent with a trickle-down theory, but this would take a longer period of study.

What evidence is there to support our approach that these differences between educational groups are a reflection of differences in tastes? The turnover in leading names varies by education of mother in ways that make no sense from the class imitation perspective (see table 6, last column). If that theory was correct, given the sequence it postulates, we would expect equally rapid turnover for each educational subgroup. Indeed, if anything, we would expect turnover to be more rapid for the higher SES since they are trying to drop any names that are imitated—

whereas the lower strata have no push to drop their choices. In point of fact, it is the lower strata who are more rapid in their name changes. (It is not immediately obvious to me how a simple imitation diffusion model would generate such a pattern.) We also find that the disposition to give children unique names is greater among less educated mothers. Since they are unique names, there is no imitation involved whatsoever. The data reported in tables 7, 8, 9, and 11 all indicate systematic differences in name choices in ways that make sense in terms of taste differences—for example, difference in the semantic differentials of the names each class chooses, the origins of the names, the sound patterns, even the likelihood that the names have a history. In my estimation, it would take an incredibly complex, convoluted, and forced interpretation to incorporate all of these results into an imitation model.

What is the evidence against this interpretation? Besnard not only says we are wrong; he also believes that our New York State data prove this and that the French data have already shown this to be the case. I decided to investigate each of these claims. In my opinion, the results are rather interesting.

I find the logic unclear in the way Besnard raises New York State names as examples of this so-called imitation process. This is because he makes the untested assumption that in New York the growth and decline in the number of children with a name is a sound indicator of class diffusion. At any rate, I take it that he claims that names in decline are names that the higher strata are dropping. The cruel truth is that the results totally fail to support him. If we compute ND (the index of net difference; see p. 527) separately for the first half and second half of the period for which we have data, here is what we find: names that are in decline—as Besnard defines it—are as likely to have a more negative ND in the first half as in the second half. Names that are expanding during this period, without exception, show lower NDs in the second half than in the first. In other words, the shifts are in the opposite direction from what I understand him to be claiming.

Where are these French studies that rigorously demonstrate that this vertical diffusion is the driving force accounting for contemporary class differences presently observed? Besnard mentions two papers written *after* our article was published, which by their titles seem to address this topic (Besnard 1994; Besnard and Grange 1993). Through the editors of the *AJS* I have requested copies of both from Besnard, but I did not receive them before writing this rejoinder. Nevertheless, I was able to find some relevant information. The latest edition of Besnard and Desplanques (1993), written after our article was published, provides a new operational definition of the phrase “adding insult to injury.” They have now added a table (pp. 282–83) that compares the leading names across

four social groups. Except for the failure to compute ND, this is beginning to look remarkably like our tables 3 and 4. Why are they doing this? A clue is provided in the previous edition where only two social categories are compared (Besnard and Grange 1991, p. 270). The commentary tells us that the social differentiation of tastes is continuing to grow; that is, there are class differences that are not simply a function of vertical imitation. This is also verified in Besnard himself (1991), where he admits that there is a growing social diversification of choices that challenges this older model. The question now becomes How important is the vertical diffusion model anyway, even in France?

There are some data for France that are hardly definitive, but they do not exactly increase one's confidence in Besnard's use of present-day France as grounds for dismissal of our approach. For fathers in each of six different socioprofessional classes, Desplanques (1986, pp. 76–79) provides data on the five names most frequently given to sons and daughters.<sup>2</sup> The most recent period is 1980–81, the earliest 1945–49. I looked at the leading names in the most recent period for the sons of either farmers or workers, asking if there is evidence that these names represent vertical diffusion from the choices made earlier by managers/executives. This was then repeated for their daughters. The results largely contradict the class imitation model. Among boys' names, for example, four of the leading names given to the sons of agriculturalists or workers in 1980–81 (Jérôme, Sébastien, Mickaël, and Cédric) *never* appear as a leading name among the sons of managers, going back to 1945–49. The concept of vertical diffusion is hardly supported by Frédéric, a leading name only among the sons in 1980–81 born to agricultural workers. This is because it first appears simultaneously among the sons of *both* managers and workers in the 1970–74 period. Likewise, Guillaume first appears as a leading name in 1980–81, simultaneously among the sons born to managers, the intermediate professions, and the agriculturalists; Nicolas first appears as a top-five name among the managers in 1970–74, then among all but the agricultural and worker classes in 1975–79, and finally among all of the classes in 1980–81. Does this look like a perfect example of vertical diffusion? Except that by 1989–90 it ranked seventh among workers and was still third among the sons born to the manager/executive group (Besnard and Desplanques 1991, p. 270)—hardly a textbook example of a vertical diffusion model that predicts that the higher strata would drop the name after the lower strata pick it up. Of eight names for boys, only Julien could be said to fit the model: it first appears as a leading name among the sons of the managerial class in 1975–79; it also appears

<sup>2</sup> Desplanques chooses not to provide information on the occupations of the mothers.

among the sons of four other classes in 1980–81, one of which is the workers. Using another source I do find that Julien is the second most popular name among workers in 1989–90, at a time when it ranks thirteenth among the managers' sons (based on Besnard and Desplanques 1991, p. 270). This is not exactly a disappearance among the elite, and the rapidity of diffusion raises questions that we have discussed earlier, but I can afford to give Besnard the benefit of the doubt because not one other name even vaguely fits this vertical diffusion model.

I will not burden the reader with a detailed examination of seven leading names that appear among the daughters born of agriculturalists or workers in 1980–81, having cited the source for these data. These also fail to support the diffusion model, with most being adopted by different classes in the same period or adopted without ever even appearing among the daughters of the manager/executive fathers. Marie seems to be a mild exception since it is adopted by agriculturalists and some other classes in the period immediately following its adoption by managers and executives for their daughters. But even here the model crashes: in 1989–90 it is not even among the top 20 among daughters born to workers, whereas it is in third place among the daughters of managers/executives. In short, if you look at the data in a systematic way, the argument against our perspective does not even hold up with the data for France in the current period.

Finally, why is Besnard's 1979 paper ignored? It consists of a confused review of the literature, two hypotheses (neither of which, to my knowledge, has he ever studied), a proposal to interview parents about their motives for choosing a name (an idea that he disowned in his 1991 paper), and a plea for quantitative research on names. Being startled to learn in his letter that these 7.5 pages of text "outlined a program and contributed to the creation of a research tradition on first names in France, to which historians, demographers and sociologists have contributed," I reread the paper in question. Try as I may, I still find no reason to cite it and cannot understand how it had such a great influence in France. Perhaps I am wrong. As a reality check, I examined the four French books Besnard cites as illustrating this tradition of research stimulated by his paper. Besnard and Desplanques (1993) does not even cite Besnard (1979)—and for good reason, considering its hypotheses; Dupâquier, Pélissier, and Rébaudo (1987) includes the paper in the bibliography, but mentions it nowhere else; only three of the 35 articles in the collection edited by Dupâquier, Bideau, and Ducreux (1984) refer to this paper (one of these being written by Besnard himself). Perouas et. al. (1984) not only dates the origins of their project to a period preceding the paper in question, but the preface (by Jean Delumeau of the Collège de France) traces the intellectual origins of the book back to the 1932 publication by Marc

Bloch and a 1941 work by Lucien Febvre—Delumeau does not even mention Besnard. Somehow this doesn't quite jibe with the way Besnard describes his role.

What about the insinuation that Besnard's paper is purposely not cited in order for our paper to claim priority on all of these discoveries? First of all, if we are going to steal Besnard's ideas, why write to him for a reprint? Also, if we are going to claim credit for originating all of the ideas in our paper, we would not have written that "the analysis of gender differences in the naming process is inspired by Alice Rossi's (1965) pioneering study of the names given to children in white middle-class families living in and near Chicago" (p. 516). Also, how can it be, as Besnard claims, that the central thesis of our paper is at one point lifted from him and, at another point, totally wrong?

I cannot understand Besnard's concern about priority; we make no such claim and—moreover—he has no business making one. The notion of studying first names to get at various social processes is a long-standing one; our paper cites several publications that long precede his, and this without even trying hard. Rossi's paper, for example, was published 14 years earlier than Besnard's. Far be it from me to write about French scholarship, but it does not appear that Besnard is the pioneer in that literature either. Even the Bloch and Febvre publications cited by Delumeau are hardly the earliest. For example, a bibliography of work on French first name studies compiled by Semeilhon (1984) lists 212 publications prior to Besnard's 1979 paper, with 33 occurring even before the 20th century. Speaking of discrepancy, the 1979 Besnard paper cites only one of the 212 French works on first names that had been published by that time.

At any rate, I completely agree with Besnard on a key point: he is not cited as extensively as he would like. The only issue is why.

STANLEY LIEBERSON

*Harvard University*

## REFERENCES

- Besnard, Philippe. 1979. "Pour une étude empirique du phénomène de mode dans la consommation des biens symboliques Le cas des prénoms." *European Journal of Sociology* 20:343–51.
- . 1991. "Le choix d'un prénom: Actualité de la méthode Durkheimienne." *Recherches Sociologiques* 22:53–60.
- . 1994. "A Durkheimian Approach to the Study of Fashion: The Sociology of First Names." Pp. 159–73 in *Debating Durkheim*, edited by W. S. F. Pickering and H. Martins. London: Routledge.

- Besnard, Philippe, and Guy Desplanques. 1991. *Un prénom pour toujours: La cote des prénoms hier, aujourd'hui et demain*. Paris: Balland.
- . 1993. *Un prénom pour toujours: La cote des prénoms en 1994*. Paris: Éditions Balland.
- Besnard, Philippe, and Cyril Grange. 1993. "La fin de la diffusion verticale des goûts? Prénoms de l'élite et du vulgum." *L'Année Sociologique* 43:269–94.
- Bloch, Marc. 1932. "Noms de personne et histoire sociale." *Annales d'Histoire Économique et Sociale* 4.
- Desplanques, Guy. 1986. "Les enfants de Michel et Martine Dupont s'appellent Nicolas et Céline." *Économie et Statistique* 184:63–83.
- Dupâquier, Jacques, Alain Bideau, and Marie-Elizabeth Ducreux, eds. 1984. *Le prénom: Mode et histoire*. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.
- Dupâquier, Jacques, Jean-Pierre Pélissier, and Danièle Rébaudo. 1987. *Le temps des Jules: Les prénoms en France au XIX<sup>e</sup> siècle*. Paris: Christian.
- Febvre, Lucien. 1941. "De la renaissance à la contre-réforme, changement de climat?" *Annales d'Histoire Sociale*.
- Perouas, Louis, Bernadette Barrière, Jean Boutier, Jean-Claude Peyronnet, and Jean Tricard. 1984. *Léonard, Marie, Jean et les autres: Les prénoms en Limousin depuis un millénaire*. Paris: Éditions du Centre National de la Recherche Scientifique.
- Rossi, Alice S. 1965. "Naming Children in Middle-Class Families." *American Sociological Review* 30:499–513.
- Semeilhon, Andrée-Geneviève. 1984. "Orientation Bibliographique." Pp. 385–95 in *Le prénom: Mode et histoire*, edited by Jacques Dupâquier, Alain Bideau, and Marie-Elizabeth Ducreux. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.

## INTERCLASS SEGREGATION, POVERTY, AND POVERTY CONCENTRATION: COMMENT ON MASSEY AND EGGERS<sup>1</sup>

In a recent article (*AJS* 95 [March 1990]: 1153–88), Douglas Massey and Mitchell Eggers provide a test of William Julius Wilson's (1987) thesis that the rising concentration of minority poverty in inner cities results largely from nonpoor minority families moving away from poor minority families. They reason that if Wilson's thesis is correct then there should be observable increases in minority interclass segregation over time and these increases in interclass segregation should explain increases in minority poverty concentration.

In this article, Massey and Eggers base their analysis on 1970 and 1980 data for 60 metropolitan areas. They measure interclass segregation with the index of dissimilarity ( $D_{xy}$ ), and they measure poverty concentration with an index of intraclass contact ( $P_i^*$ ) computed for the poor (class isolation);  $D_{xy}$  measures the degree of spatial separation of social groups. It is a symmetrical index that gives the proportion of  $X$  and  $Y$  members

<sup>1</sup> Direct all correspondence to Craig St. John, Department of Sociology, University of Oklahoma, Norman, Oklahoma 73019. Many thanks to Sampson Blair for assistance with the graphics.

that would have to exchange residential locations to achieve an even distribution. Among members of the same social group  $xP_x^*$  measures the likelihood of residential contact. It gives the probability that a member of group  $X$  will encounter another member of group  $X$  by virtue of sharing the same residential location. In this comment I focus on Massey and Eggers's analysis for blacks.

From examining the 1970–80 change in interclass segregation for blacks, Massey and Eggers conclude that increases in interclass segregation among blacks in the 1970s were not large enough to account for significant increases in black poverty concentration during the decade. This judgment is confirmed in a multivariate analysis in which, for the period 1970–80, change in the isolation of poor blacks is regressed on change in interclass segregation among blacks, change in black residential segregation from whites, and change in the black poverty rate. In this regression, change in interclass segregation does not have a statistically significant effect on change in class isolation. Massey and Eggers interpret these results as evidence that Wilson's thesis is incorrect. Further, these results are important for an additional *AJS* article (Massey 1990) and a book (Massey and Denton 1993) in which it is argued that racial residential segregation, not the selective out-migration of nonpoor blacks, is responsible for making the underclass

Wilson (1991) has criticized the results of this study on two grounds. First, he argues that using an index of class isolation provides for a description of the overall level of concentrated poverty in a metropolitan area, but neither identifies specific neighborhoods that are ghettos nor measures increases in the concentration of poor blacks living in such neighborhoods. Second, he argues that change in the interclass segregation index is a poor measure of the extent to which the minority nonpoor have moved away from the minority poor, because it does not detect much of this movement. As many of the nonpoor move away from the poor, they end up in new neighborhoods with smaller but substantial poor populations. This movement does not necessarily contribute to an increase in interclass segregation, because, as the poor in the old neighborhoods have become more segregated from the nonpoor, those in the new neighborhoods have become less segregated.

In this comment I argue that there are two additional reasons why change in the interclass segregation index does not detect the movement of nonpoor blacks away from poor blacks. The first concerns the way Massey and Eggers calculate the interclass segregation index. They divide the black population into four classes: the poor (P), the lower middle class (LM), the upper middle class (UM), and the affluent (A). They compute interclass segregation by averaging the six segregation indexes calculated for the six combinations of these classes taken two at a time

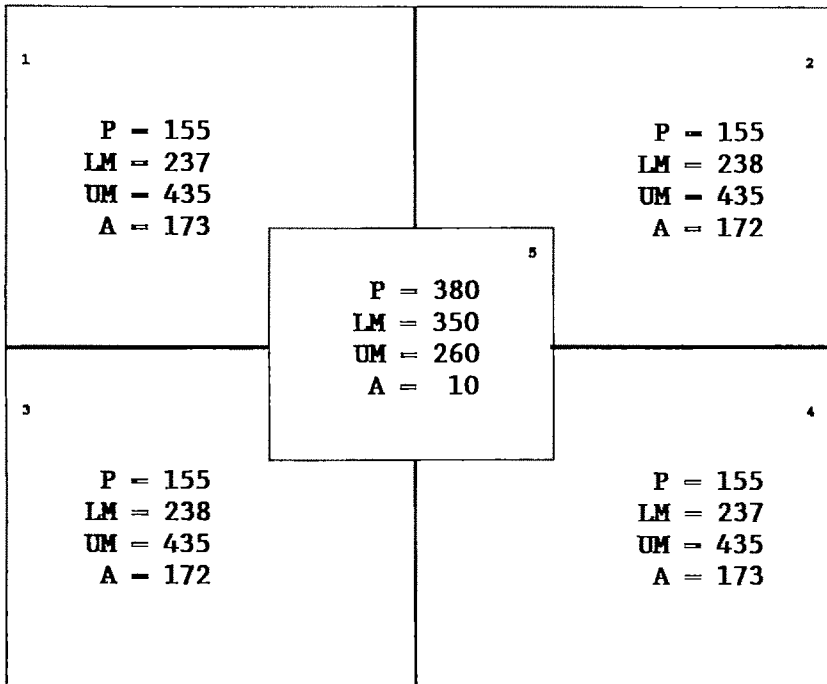


FIG. 1.—Simulation for average interclass segregation index, baseline condition:  $D_{P,LM} = .11$ ,  $D_{P,UM} = .25$ ,  $D_{P,A} = .37$ ,  $D_{LM,UM} = .14$ ,  $D_{LM,A} = .26$ ,  $D_{UM,A} = .12$ , average  $D = .21$ ,  $pP^* = .24$ .

( $D_{P,LM}$ ,  $D_{P,UM}$ ,  $D_{P,A}$ ,  $D_{LM,UM}$ ,  $D_{LM,A}$ , and  $D_{UM,A}$ ). If lower-middle-class, upper-middle-class, and affluent blacks move away from poor blacks, as Wilson argues they have, then the three segregation indexes involving poor blacks and other blacks ( $D_{P,LM}$ ,  $D_{P,UM}$ , and  $D_{P,A}$ ) will increase, increasing the average segregation index. Simultaneously, the three segregation indexes not involving poor blacks ( $D_{LM,UM}$ ,  $D_{LM,A}$ , and  $D_{UM,A}$ ) are likely to decrease, decreasing the average segregation index. To the extent these countervailing effects cancel each other out, it is possible for poor blacks to become more segregated from nonpoor blacks, as the latter move away from the former, while there is no change in average interclass segregation.

This situation is illustrated in the simulation presented in figures 1 and 2. Consider the city in figure 1 with five neighborhoods and 5,000 people. Of these, 1,000 are poor, 1,300 are lower middle class, 2,000 are upper middle class, and 700 are affluent. This distribution approximates the distribution Massey and Eggers report for blacks in the Chicago standard metropolitan statistical area (SMSA) for 1970. These people are distrib-



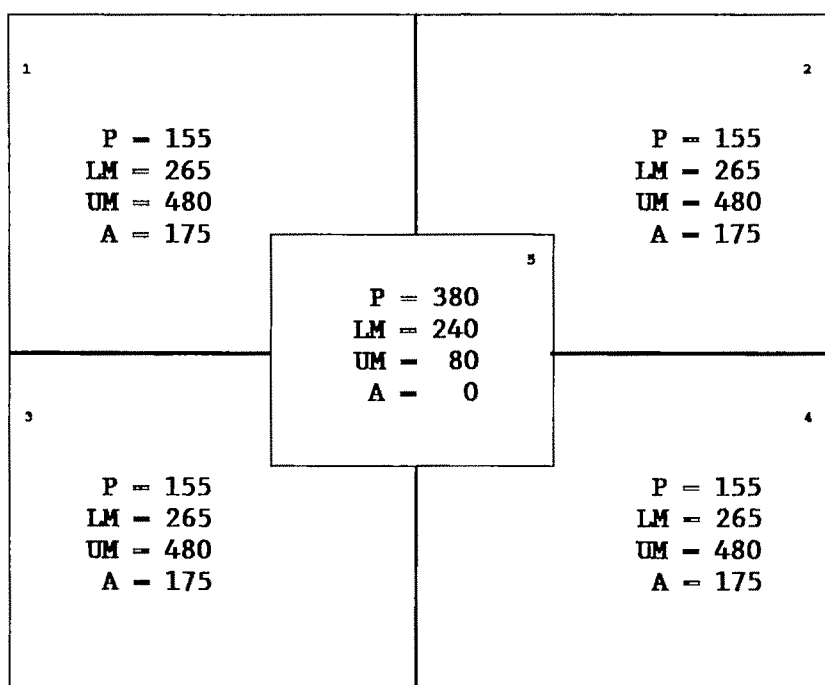


FIG. 2.—Simulation for average interclass segregation index, effect of out-migration of nonpoor:  $D_{P,LM} = .20$ ,  $D_{P,UM} = .34$ ,  $D_{P,A} = .38$ ,  $D_{LM,UM} = .14$ ,  $D_{LM,A} = .18$ ,  $D_{UM,A} = .04$ , average  $D = .21$ ,  $rP^2 = .30$ .

uted into the five neighborhoods such that there is one neighborhood approximating a ghetto with a poverty rate of 38% surrounded by four mixed-income neighborhoods with poverty rates of 16%. The index of class isolation of the poor is .24. The six interclass segregation indexes for the city are as shown and the average is .21. In figure 2, 110 lower-middle-class blacks, 180 upper-middle-class blacks, and 10 affluent blacks (300 total) have moved from neighborhood 5 evenly into neighborhoods 1–4. Neighborhood 5 now has a poverty rate of 54%, and the other neighborhoods now have poverty rates of 14%. The index of class isolation of the poor is now .30. The six new interclass segregation indexes are as shown. Poor blacks became more segregated from nonpoor blacks, but the levels of segregation among the nonpoor blacks decreased, canceling out the increased segregation of poor blacks. The average interclass segregation index for the city remains .21. Thus, in spite of the substantial selective out-migration from neighborhood 5 that significantly increased the isolation of the poor, the average interclass segregation index did not change at all.

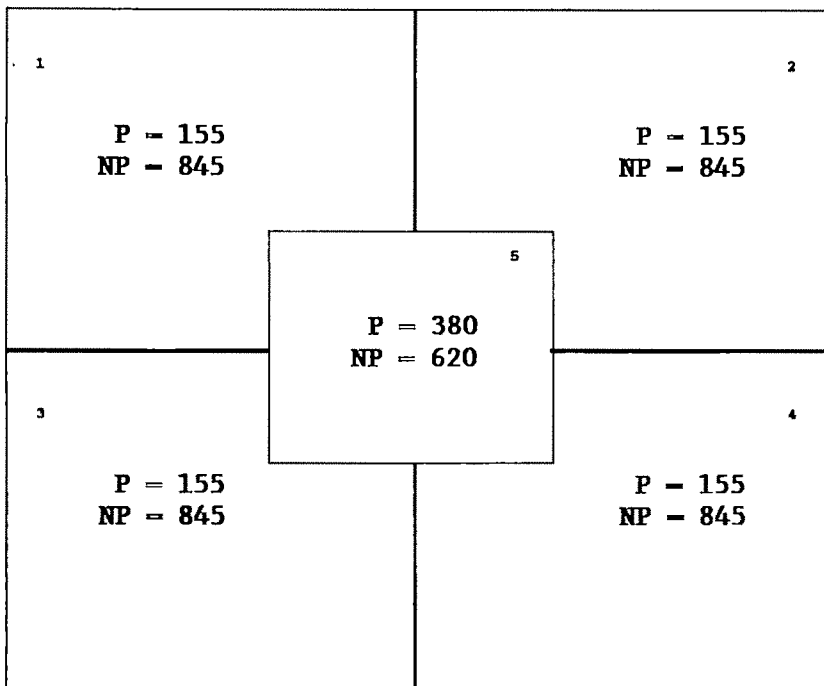


FIG. 3.—Simulation of out-migration of nonpoor and increasing poverty rate, baseline condition:  $D_{P,NP} = .23$ ,  $P_{NP}^* = .24$ .

The second reason why change in the interclass segregation index does not detect the movement of nonpoor blacks away from poor blacks involves the increasing poverty rates in the metropolitan areas in Massey and Eggers' study between 1970 and 1980. Increasing poverty in these metropolitan areas simultaneously served to decrease black interclass segregation and increase the isolation of poor blacks. Decreases in interclass segregation caused by increasing poverty offset increases in interclass segregation that might result from the nonpoor moving away from the poor. Thus, increasing poverty can make it appear as if the nonpoor have not moved away from the poor, even when they have, while also contributing to the increasing isolation of the poor.

This situation is illustrated in the simulation presented in figures 3–6. Consider the city in figure 3. This figure is the same as figure 1 except lower-middle-class, upper-middle-class, and affluent blacks are included in one class, the nonpoor, for ease of presentation. There are 1,000 poor people (P) and 4,000 nonpoor people (NP) in this city for an overall poverty rate of 20%. The index of interclass segregation ( $D_{P,NP}$ ) is .23,

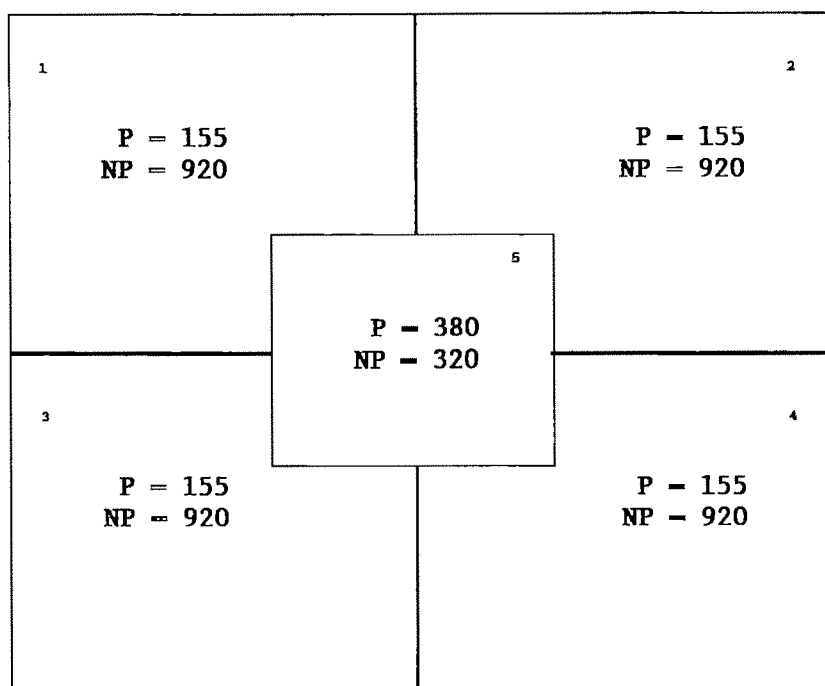


FIG. 4.—Simulation of out-migration of nonpoor and increasing poverty rate, effect of out-migration of nonpoor:  $D_{P,NP} = .30$ ,  $P_{NP}^* = .30$ .

and the index of class isolation for the poor ( $P_{P}^*$ ) is .24. Now consider figure 4 after 300 of the nonpoor have moved from neighborhood 5 evenly into neighborhoods 1–4. The overall poverty rate is still 20% although the poverty rate has increased substantially in neighborhood 5 and decreased slightly in neighborhoods 1–4. The index of interclass segregation is .30, an increase of .07, and the index of class isolation for the poor is .30, an increase of .06. The increase in interclass segregation is a direct result of the movement of the nonpoor from neighborhood 5, and it corresponds almost exactly with the increase in the isolation of the poor. The comparison of figures 3 and 4 corresponds with Wilson's thesis.

Figure 5 is the result of taking the city shown in figure 3 and increasing the poverty rate by 10 percentage points in each neighborhood with no movement of the nonpoor away from neighborhood 5. Compared to figure 3, the economic composition in each neighborhood now more closely matches the economic composition of the city as a whole, and the poverty rate in each neighborhood in which poor people live is higher. Thus, the index of interclass segregation for the city in figure 5 is .17, a decrease of .06 from figure 3, and the index of class isolation for the poor is .33,

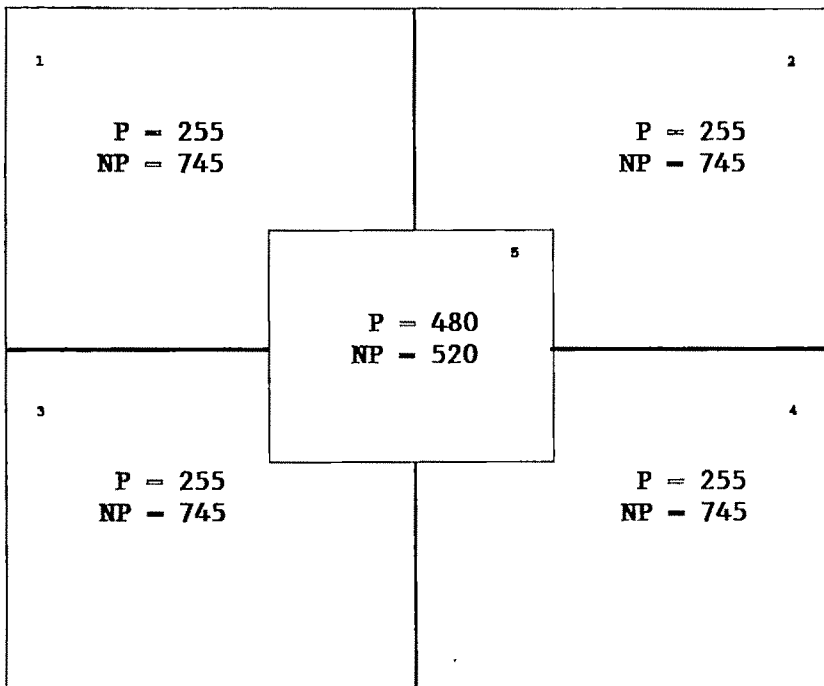


FIG. 5.—Simulation of out-migration of nonpoor and increasing poverty rate, effect of increasing poverty rate:  $D_{P,NP} = .17$ ,  $P_{NP}^* = .33$ .

an increase of .09 from figure 3. Increasing poverty has both reduced interclass segregation and increased the isolation of the poor.

Figure 6 displays the combined effects of increasing poverty and movement of the nonpoor away from the poor. The distribution of the poor and the nonpoor into neighborhoods in this panel is the result of first increasing the poverty rate in every neighborhood of the city in figure 3 by 10 percentage points and then having 300 of the nonpoor move from neighborhood 5. The interclass segregation index is .26, an increase of only .03 compared to figure 3, and the index of class isolation for the poor is .38, an increase of .14 compared to figure 3. Because the positive effect of the movement of the nonpoor from neighborhood 5 on interclass segregation has been offset by the negative effect of increasing poverty, there has been little increase in interclass segregation. At the same time, both increasing poverty and the movement of the nonpoor from neighborhood 5 have contributed to the increase in the index of class isolation for the poor.

Consequently, for the two reasons I have specified, finding that there is little change in interclass segregation while there are increases in class isolation does not mean that nonpoor blacks have not moved away from

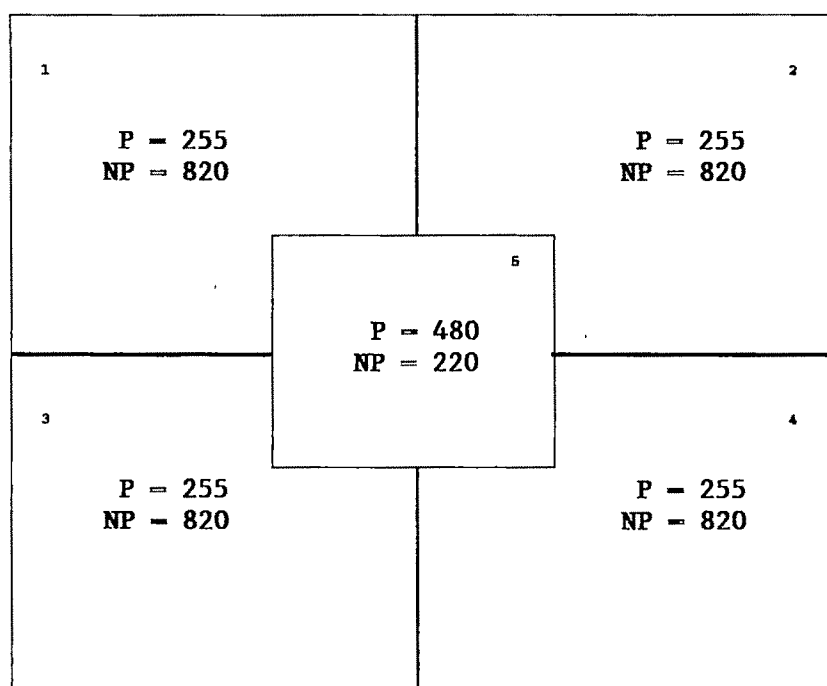


FIG. 6.—Simulation of out-migration of nonpoor and increasing poverty rate, combined effect of out-migration of nonpoor and increasing poverty rate:  $D_{P,NP} = .26$ ,  $P_{NP}^* = .38$ .

poor blacks nor that this movement was inconsequential for the increasing isolation of poor blacks. On the contrary, it is almost certain that changes in measures of overall interclass segregation calculated for metropolitan areas obscure population dynamics at the neighborhood level. To understand the role of selective out-migration of nonpoor blacks in creating the underclass requires an examination of population gains and losses at the neighborhood level and the subsequent changes in neighborhood social class composition.

CRAIG ST. JOHN

*University of Oklahoma*

#### REFERENCES

- Massey, Douglas S. 1990. "American Apartheid: Segregation and the Making of the Underclass." *American Journal of Sociology* 96:329–57.  
 Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation*

- and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Massey, Douglas S., and Mitchell L. Eggers. 1990. "The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970-1980." *American Journal of Sociology* 95:1153-88.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- . 1991. "Studying Inner-City Social Dislocations: The Challenge of Public Agenda Research." *American Sociological Review* 56:1-14.

## BEYOND THE TECHNICAL DETAILS: REPLY TO ST. JOHN

In our article, we computed measures of black interclass segregation to examine Wilson's hypothesis that the geographic concentration of black poverty stemmed from the out-migration of nonpoor blacks from poor ghetto neighborhoods (Massey and Eggers 1990). We demonstrated that increases in income segregation among blacks were small during the 1970s and were unrelated to trends in the concentration of black poverty. For two reasons, St. John argues that this approach might not detect the movement of nonpoor blacks away from poor black neighborhoods.

First, he shows that under certain circumstances nonpoor out-migration can produce increases in some measures of interclass segregation and decreases in others, so that when average indexes are computed the changes offset one another to yield little change in the summary measure of class segregation that we used. We agree that this outcome is possible, depending on where nonpoor out-migrants from poor black neighborhoods are assumed to settle. This argument, however, does not explain why we find the same results using the affluent-poor segregation index, where there are no offsetting effects.

Second, St. John shows that under certain circumstances an increase in poverty among blacks can offset the effect of nonpoor blacks moving out, again leading to little change in our summary measure of interclass segregation. We also agree this scenario is possible, although it requires poverty rates to increase in neighborhoods where nonpoor blacks have relocated. This assumption is probably not realistic, since nonpoor movers tend to move toward higher status neighborhoods, which are less likely to bear the brunt of any increase in poverty.

In general, we concur with St. John's main point that the results we reported in "The Ecology of Inequality" cannot definitively reject Wilson's black middle-class migration hypothesis. Our findings simply add to a growing list of circumstantial evidence that is inconsistent with this view.

If black poverty concentration is caused by the out-migration of non-poor blacks from poor neighborhoods, why are levels of poverty concentration so much lower for Asians and Hispanics though their levels of income segregation are so much higher, as we reported? And why is the level of black interclass segregation in different metropolitan areas unrelated to the level of net out-migration from poor black neighborhoods, as reported in Massey, Eggers, and Denton (1994)? And why is the poverty rate across census tracts unrelated to the level of net out-migration from those tracts, as reported in Massey and Kanaiaupuni (1993)? And if nonpoor blacks are moving out of ghetto areas, where are they going? Certainly not to white areas: In metropolitan areas that experienced a sharp increase in black poverty concentration, levels of racial segregation hardly changed (Massey and Denton 1987) and black-white segregation showed no tendency to decline as income rose (Denton and Massey 1988).

The problem with all of this evidence, of course, is that it is indirect. None of it directly measures the movement of individual blacks, poor and nonpoor, in and out of specific kinds of neighborhoods. This analysis is provided elsewhere (Massey, Gross, and Shibuya 1994), and we hope that it persuades policymakers and scholars to reject the idea that non-poor blacks somehow bring misery to their former neighborhoods by moving away in an attempt to improve their own residential circumstances. We view this argument as a specious attempt to blame the victims of prejudice and discrimination for the consequences of their own victimization.

Why should the black middle class be different from the middle class of other groups? Throughout American history, upwardly mobile individuals and families have tried to improve their welfare and their opportunities by moving to better neighborhoods, but racial segregation has closed this avenue of advancement to aspiring African-Americans.

As a result, middle-class blacks have been less able to achieve improved neighborhood conditions than middle class members of other groups, and poor blacks have been forced to endure concentrations of poverty that are unparalleled. Rather than blaming middle-class blacks for the emergence of the urban underclass, scholars and policymakers should work to open up housing markets and give African-Americans access to one of the most important avenues of upward socioeconomic mobility the country has to offer.

DOUGLAS S. MASSEY

*University of Pennsylvania*

MITCHELL L. EGGERS

*Economic Commission for Europe*

REFERENCES

- Denton, Nancy A., and Douglas S. Massey. 1988. "Residential Segregation of Blacks, Hispanics, and Asians by Socioeconomic Status and Generation." *Social Science Quarterly* 69:797-817.
- Massey, Douglas S., and Nancy A. Denton. 1987. "Trends in the Residential Segregation of Blacks, Hispanics, and Asians." *American Sociological Review* 52:802-25.
- Massey, Douglas S., and Mitchell L. Eggers. 1990. "The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970-1980." *American Journal of Sociology* 95:1153-88.
- Massey, Douglas S., Mitchell L. Eggers, and Nancy A. Denton. 1994. "Disentangling the Causes of Concentrated Urban Poverty." *International Journal of Group Tensions* 24:267-316.
- Massey, Douglas S., Andrew B. Gross, and Kumiko Shibuya. 1994. "Migration, Segregation, and the Spatial Concentration of Poverty." *American Sociological Review* 59:425-45.
- Massey, Douglas S., and Shawn M. Kanaiaupuni. 1993. "Public Housing and the Concentration of Poverty." *Social Science Quarterly* 74:109-23.



## Book Reviews

*Paradoxes of Gender.* By Judith Lorber. New Haven, Conn.: Yale University Press, 1994. Pp. xi + 424. \$30.00.

Rosanna Hertz  
Wellesley College

Over the past two decades, gender has moved from being a marginal subfield of family sociology to holding a central position in the discipline. Judith Lorber has played a vital role in the transformation of gender studies, and this book reflects her extensive influence over the way sociologists understand gender. Because of work like hers, the study of gender is (and will continue to be) taken seriously.

*Paradoxes of Gender* offers not only a synthesis of the state of the field, but, in contrast to earlier handbooks and critiques of gender dichotomies, it is framed in a clear and sophisticated context that cuts across levels of analysis. At the very outset Lorber makes clear that we must view gender as a social institution that "establishes patterns of expectations for individuals, orders the social processes of everyday life, is built into the major social organizations of society, such as the economy, ideology, the family, and politics, and is also an entity in and of itself" (p. 1). By emphasizing that gender is a "human invention" (p. 6), she leaves no doubt that hers is a sociological perspective and one that does not easily or comfortably accept the notion that there is anything "natural" (p. 284) about gender inequality.

One of the greatest strengths of this book is that it is packed with new insights from research that may be familiar to many scholars. The first third of the book challenges the myth that gender is a simple biological construct. According to Lorber, the social construction of gender begins when an infant is assigned to a sex category on the basis of genitalia. Individuals assigned to one sex category are treated differently from individuals assigned to the other. In other words, physical anatomy does not fix a person's gender; however, once gender is assigned, individuals are expected to adhere to appropriate norms and values for women and men.

Moreover, Lorber contends, gender is not a simple dichotomous variable. Using non-Western studies and historical material, she argues that these categories are much more fluid and bendable. Some non-Western societies have third and fourth genders that link genitalia, sexual orientation, and gender status in different ways (e.g., Native American berdache, Hijras in India). In this regard, gender status (living life as a man

Permission to reprint a book review printed in this section may be obtained only from the author.

or a woman) may be distinct from gender identity (being a member of a group, women or men) and from sexual desire for a woman or man. Further, in Western cultures we continue to conflate sex, sexuality, and gender, even though there are many examples that suggest that these categories do not completely overlap. For example, transvestites, bisexuals, gay men, and lesbians may cross the boundaries between sex, sexuality, and gender—demonstrating how socially constructed gender really is—but paradoxically do not disturb how deeply the social world is gendered.

The second part of the book examines gender as a social institution and how human beings have used gender to organize social life. She reviews the early emergence of inequality between men and women in foraging groups and hunting and gathering societies. The early unequal exchanges between women and men that emerged out of kinship ties continue to persist in contemporary societies. She demonstrates that paradoxes that emerged at that period remain relevant today. For instance, if men need women more than women need men, why do women end up controlled and subordinated? Similarly, in her discussion of modern-day life, including the role that parenting and children play in subordinating women, she incorporates the literature on infertility in order to understand what we can learn about gender relations when parenthood becomes difficult to achieve in traditional ways. She also provides a more holistic assessment of “traditional” claims of motherhood by referring to studies of how women of different social classes, racial groups, and sexual orientations experience motherhood. Finally, in reviewing the literature on occupational segregation, she emphasizes what changes do and do not take place when women enter male-dominated occupations and when men enter female-dominated occupations.

The last section of the book focuses on how the system of gender inequality might be dismantled. Ultimately, Lorber calls for a kind of affirmative action policy where equal numbers of girls and boys would be educated and trained for each occupation, where half the elected positions would go to men and to women, where men and women would equally parent, and so forth. This more balanced world, she contends, would lead to radical shifts in every field as women’s (and men’s) views and experience would now alter how knowledge is produced, taught, and interpreted and how social problems are defined and deemed worthy of pursuit. This utopian vision may lead to gender equality; it is unclear to me, however, that it will eradicate inequality more generally. In all likelihood, it will just shift the focus of inequality. In this respect—and despite her claim that the analysis of gender as a social institution represents a paradigmatic shift—the *fact* of gender remains deeply rooted in virtually every aspect of social life.

A real strength of the book is the way Lorber weaves together several different literatures in new ways. Throughout I was struck by two aspects of Lorber’s approach: first, how skillfully she uses the exceptional cases and experiences to inform what we know and can learn about gender

and, second, how powerfully she establishes the concept of paradox in analyzing divisions between women and men and their resulting inequalities.

*Paradoxes of Gender* deserves to be read by both specialists and general readers alike. Those who are familiar with the literature will find Lorber's insights and her ability to reorganize various pieces of the puzzle of gender construction unmatched. Those scholars who want an introduction to gender studies will find her book to be an intelligent and thoughtful summary.

*A Very Social Time: Crafting Community in Antebellum New England.* By Karen V. Hansen. Berkeley and Los Angeles: University of California Press, 1994. Pp. xv + 262. \$30.00.

Jeanne Boydston  
*University of Wisconsin—Madison*

Keenly aware of the many missteps possible in crossing into unfamiliar fields, scholars tend to shy away from interdisciplinary work. Karen V. Hansen's *A Very Social Time: Crafting Community in Antebellum New England* is a good reminder that such journeys, often necessary, can be gracefully conducted.

*A Very Social Time* arises from Hansen's skepticism about the adequacy of a two-sphere model for contemporary social science research on gender. Hansen, a sociologist, decided to investigate the problem at its historical roots—in the daily life of antebellum New England, where Americans first fully conventionalized the ideology of separate gender spheres. Poring through 56 diaries, 20 letter collections, and 19 autobiographies, by both men and women, she found instance after instance of the broad social interdependence of men and women—shared activities, shared space, shared customs, and a range of shared mechanisms through which women and men alike approved or censured the behavior of other men and women in their community. "These vital accounts," she concludes, "not only challenge the way contemporary scholars frame social structure but also call into question the general acceptance of a broad, dominant middle-class culture and the supposition that nineteenth-century women led lives entirely separate from men" (p. 1).

Although this is not a new conclusion (historians have long recognized that few real people enacted literally the highly stylized gender feints of the ideology of spheres), Hansen is to be applauded for her delightfully fresh rendering of daily life in mid-19th-century New England. Armed with a full command of the historical literature and an impressive capacity to read primary documents with care and intelligence, she offers a deft evocation of the clamorous, busy, and yet highly personalized sociability of this lost world.

Hansen's aim is not merely to note the historical anomalies of the two-sphere model, however, but to suggest a revision. To the conventional "public" and "private" she adds "a third, mediating category" (p. 7), the "social," which "embraces activities that transcend individual households and operate independent of state, such as visiting, gossiping, churchgoing, attending lectures, joining political movements, baby-sitting a neighbor's child, and shopping" (p. 9). "In antebellum New England," she argues, "the social was characterized by rules and negotiation, as compared to the laws and litigation of the public, and in contrast to the motivating forces of emotion and need in the private. The social's unique values included mutuality, reciprocity, voluntarism, and localism" (p. 9).

Hansen is right about the need for a new framework. Whether this book completely answers that need is another question. She offers the social as a third sphere, alongside but not in place of the public/private dichotomy, which she wishes to retain since it "continues to hold an unyielding grip on both the popular imagination and academic discourse" (p. 7). I agree with the observation but demur on the significance of longevity: It would not be the first time a bad idea had held the world spellbound.

Hansen defines the social in relation to the individual household and the state, a maneuver requiring that she posit relatively stable identities for both of these poles. But were they stable? (Are they now?) Certainly, "household" was not: Hundreds of thousands of migrants pulled up stakes to travel west; sons and daughters bailed out of natal families in great numbers; native-born families coupled intense intimacies with a willingness to hire perfect strangers to work in their homes; and workers resided in apartments two or three families to the room. Meanwhile, the antebellum state stumbled along as best it could—an ungainly collection of competing interests, unruly behaviors, and underdeveloped agencies.

The ideas of "household" and "state" operating in this study, then, are themselves somewhat idealized. Those idealizations stripped away, the qualities Hansen finds unique to the "social"—especially those qualities that belie a certain contingency and absence of formal structure—appear to characterize most aspects of this extremely volatile period. The "social," it seems to me, is less the third partner of "public" and "private" than their true nature revealed.

One implication of *A Very Social Time* is that we must look beyond the icons of "public" and "private" to other qualities of modern society that they obscure. Hansen emphasizes mutuality, reciprocity, voluntarism, and localism. My own reading of antebellum history suggests that those qualities coexisted with, and were often marshaled in the service of, opportunism, unequal dependency, exclusion, and hierarchy. What the icons have concealed, I suspect, is not only reciprocity but ruthlessness. The question of why that concealment grounded its authority in the language of gender is left unexplored in this study.

*Encounters with Aging: Mythologies of Menopause in Japan and North America.* By Margaret Lock. Berkeley and Los Angeles: University of California Press, 1993. Pp. xliv + 439.

L. L. Cornell  
*Indiana University*

*Encounters with Aging* is a masterwork of thought about gender, the body through time, and cultural difference between Japan and North America. It is rich in detail, for its narratives of the lives of middle-aged women are long and textured. They will bear careful study. More importantly, the book is rich in thought about a portion of human life that has not been thought much about: the middle years.

This book originates from one observation: women stop menstruating. But rather than following the American self-help book idiom and bemoaning fearsome physical changes, it uses this observation as a gateway to a wide variety of issues. Language is one. The characteristic symptom of menopause for European-origin women in Canada and the United States is hot flashes. Hot flashes are so uncommon in Japan (only 9% of Lock's survey sample of 1,100 menopausal women report experiencing this symptom in the previous two weeks, vs. 30% in Manitoba and Massachusetts samples) that there is not even a word for the phenomenon. The hormonal change may be the same, but its physical manifestation is different.

The invention of menopause as a deficiency disease, to be treated by estrogen replacement therapy, is a second theme. It happened in the 1930s with the isolation of estrogen and progesterone; until then menopause was normal, natural, and invigorating. Western women now can avoid the atrophy of the breasts, the decrepit ovaries, the deranged endocrine system that clinicians observe and the "miserable and emotionally disturbed time" that Freudian therapy promises (p. 329).

The many Japanese doctors whose words Lock reproduces here are astonished and horrified by this regime. Few women come to them with symptoms of menopause; those who do visit doctors report not hot flashes but shoulder stiffness and headaches. Doctors may use hormones for a short time in the most severe cases, but they prefer to treat these complaints with herbal medicine from the Chinese tradition and to advise the women presumably at greatest risk, pampered housewives with nothing to do, to find some purpose for their lives. This reluctance to use hormones, in a society which is well educated and otherwise highly medicated, reflects a different epidemiology (elderly women become bedridden by stroke, not broken hips), an emphasis on mild doses of all drugs, and a fear of side effects that makes birth control pills unavailable to Japanese women (54% of Lock's broad socioeconomic sample report having one or more abortions [p. 262]).

In fact, Japanese women do not have menopause; they have *konenki* instead. *Konenki* is like Western menopause in being a socially con-

structed experience, but it differs in that it is a period of time, not a single event that divides the lifetime into "youth" and "death." Stopping menstruating is not even mentioned by many of Lock's narrators. *Konenki* is a time of changing roles: Japanese women emphasize "the way in which in middle age a woman turns from being concerned primarily with children and their care to enjoy a brief spell of freedom ('mother's time of rebellion'), before she becomes fully occupied with the care of aged people for a good number of years" (p. 45). Even the divorced, widowed, never married, and childless women Lock takes care to include in the sample define themselves as persons engaged by family responsibilities in a way unfamiliar to their more self-centered peers in North America.

Lock's method combines survey research with narratives to depict the subjective experience of ordinary women, not the illness experience of those who make their way into clinical samples. As an anthropologist she justifies the use of statistics more than most sociologists would. I believe her picture of the normative role statistics play in creating the image of the ideal Japanese family is overwrought, though it is true that statistics are ubiquitous, while access to the underlying data is surprisingly restricted.

*Encounters with Aging* will be especially valuable to students of the sociology of medicine, the life course and aging, gender studies, and comparative family sociology. It contains excellent vignettes of the details of women's lives in Japan: a woman caring for her incontinent mother-in-law while her husband has an affair (pp. 281-86); the single woman, continuously employed full-time, who makes one-third of the salary her male peers do, and thus cannot afford her own house (p. 175); the wife, a single mother by virtue of her husband's transfer to a different city, who takes up volunteer work to relieve her loneliness and headaches, but surrenders when his company forbids it (pp. 291-92).

Overall, the book gives a troubling picture of the misogyny both societies offer to aging women. Japanese worry that the prosperous and egalitarian nuclear family will obliterate the self-sacrifice of middle-aged women that assures family rather than government care for the elderly. North Americans have a far more sinister approach: a desperate clinging to the trappings of youth that belies a pathological fear of aging and death. *Encounters with Aging* challenges both views.

*Masculinity and the British Organization Man since 1945*. By Michael Roper. New York: Oxford University Press, 1994. Pp. xi + 259. \$37.00.

Mary Blair-Loy  
University of Chicago

In *Masculinity and the British Organization Man since 1945*, Michael Roper explores the identity and gender attitudes of postwar British career

managers. He interviews 25 men and five women who began their careers in the manufacturing sector of postwar Britain and who reached their late careers in the 1980s. Roper elicits intimate accounts of the managers' life histories, and he provides a subtle interpretation of his data.

In contrast to descriptions of organization men as rational, detached decision makers, Roper finds organization men living in an emotional economy in which they "negotiate the satisfactions and discontents of masculinity" (p. 2). This emotional economy helps explain the managers' passionate commitment to their work and the systematic exclusion of women from senior management.

Roper finds that relations between men at work are permeated with an intensity otherwise reserved for family and erotic relationships. Organization men establish bonds with their mentors that can take the form of dutiful son and good father, parricidal son and bad father, or young virgin and experienced seducer. Relations between peers are often charged with homoerotic affection. And organization men take pleasure in their company's products, regarding them as progeny or as objects of desire.

Roper argues that organization men from both middle- and working-class backgrounds feel their masculinity to be under attack. This sense of besiegement is in part due to the cross-cutting of gender and class hierarchies. In the cult of toughness, working-class manufacturing jobs are "hard" and productive and constitute proof of the worker's masculinity. Managers of middle-class origins privately fear that their university backgrounds and "soft" desk jobs do not warrant the same masculine pride as working-class jobs. Managers of working-class backgrounds can boast of having once done "tough" jobs, yet they feel insecure in their leadership abilities and social standing in the British class system.

The decline of the manufacturing sector further undermines the masculinity of organization men. Their power is also threatened by the ascendance of a new generation of business-school-trained managers, who emphasize company profits rather than products. Roper claims that threats from a changing economy and a new breed of managers intensify organization men's identification with company products and technology.

Roper argues that this precarious sense of masculinity prompts organization men to divide the world into masculine and feminine spheres and to reject what is culturally considered feminine, including the tender sides of their own personalities. Their siege mentality makes them feel easily threatened by strong women. More than one of his respondents had fired capable female managers.

It is impressive that Roper transcends the type of gender perspective that reduces masculinity to an unchanging phenomenon that affects all men similarly. The book focuses on one generation of managers in one nation and in one industry. The organization men's self-images and actions are affected by the national economic context and by their position in history between the family capitalists and the business-school-trained

managers. This historical grounding helps Roper avoid treating masculinity in a categorical way.

The book's strength is at the same time its weakness. Roper's focus on one generation makes it difficult for him to marshal evidence for his statements that these organization men are different from the managers that preceded and followed them. For example, he amply documents organization men's fetish for company products but presents less convincing evidence of how this resembled or contrasted with the feelings other generations of managers had toward their products. Further, although his data support his theory that organization men exclude capable women in shoring up an embattled masculinity, we are left with little insight into why family capitalists barred women from managerial careers and why the new business-school generation continues to resist the inclusion of women in senior management.

While Roper corrects the tendency of most sociologists to ignore managers' emotions and sexuality, he sometimes goes overboard in seeing gender significance almost everywhere. He tends to overdetermine organization men's beleaguered masculinity and to see it as influencing almost every statement. For example, he interprets working-class managers' confessions that they hide their fear of defeat as evidence of the masculine cult of toughness (pp. 115–16). Instead, one could see these statements as expressions of the challenges of leadership, whether faced by men or women. Other statements that he interprets as indicating erotic bonds between managers could be read as simply expressing human warmth (e.g., p. 137).

Nonetheless, Roper's book remains an impressive piece of research. His interviewing skills elicited sensitive and illuminating data. He grounds his nuanced interpretation of life histories in British economic history. The book's emphasis on the cultural and psychological dimensions of gender inequality will influence future research on gender and organizations.

*Cinderella Goes to Market: Citizenship, Gender, and Women's Movements in East Central Europe.* By Barbara Einhorn. London: Verso, 1993. Pp. viii + 280. \$59.95 (cloth); \$17.95 (paper).

Barbara Heyns  
New York University

Amidst a burgeoning literature on gender and economic transition, this book is the first to my knowledge to survey the full gamut of change in east central Europe from a feminist perspective, with the aim of making theoretical and comparative sense of postcommunism. Edited collections that offer fascinating country-specific accounts and serve as a forum for women from the region are more typical. In contrast, this book tackles



the big questions head-on, in detail, and with a wealth of comparative information. It succeeds in presenting both a cogent summary and informed, balanced interpretations. Einhorn is well versed in the theoretical issues at stake; she draws on historical and contemporary sources, on personal interviews, and on literary as well as social science material to illustrate her points. But what makes the book truly exciting is the rich set of theoretical ideas that shape the discussion.

Did socialism aim to emancipate women but simply fail? Or was the promise of gender equality only rhetoric and illusion? Will the "woman question" always be the last, unresolved issue for citizenship rights? Why are women now losing the jobs, political positions, and economic status they earned under socialism? Is there an internal logic to emerging discourses—whether nationalist or liberal—that inherently disenfranchises women relative to men? And why are Eastern European women so "allergic to feminism," while seemingly unable to mobilize in their own interest? Starkly framed, these questions embody a certain Western bias; they have, however, become fundamental to the East-West dialogue. Einhorn addresses them with care, insight, and thoughtfulness; the result is a book that must be taken seriously in all future work.

By citizenship, Einhorn means more than democratic elections or social entitlements. Rather, empowerment and the capacity for agency within a transformed civil society are central, ideals that were common among the former opposition movements. It is not yet clear to what extent the new postcommunist governments will enable citizens, be they male or female, to realize such goals; apparently, the heady attainment of political sovereignty comes first. With vivid examples, Einhorn presents the political and economic dilemmas shared throughout the region, while being sensitive to the distinctions among countries. One could fault the book for premature crisis mongering; I believe, however, that Einhorn is prescient. Pulling the plug on state socialism, at least at the outset, appears to mean losing the bath, the baby, and much of the plumbing.

The public/private distinction acquired a different meaning under state socialism than in the West. Private life was a political category, to be sure, but the concept included everything opposed to state control. For both men and women, family life and informal social networks represented a sanctuary from the state and a major source of personal freedom and autonomy. The private sphere became a base for political opposition and an idealized alternative to the command economy. Family and friends were extremely important, both objectively and subjectively, for acquiring consumer goods, for sharing information, and for simply surviving. Women "manned" the household and network nodes that kept this system functioning. The fact that such relationships are now less necessary diminishes women's influence. Privatization has become a gendered concept. For men it means entrepreneurship, for women domesticity. Understanding how these meanings evolved, however, helps to explain why many women romanticize home life without paid employment, even or especially when this is not a viable option.

The "allergy" to Western-style feminism is widespread and difficult to understand. State socialism rendered women's issues obsolete or part of a utopian future that could be implemented only under communism. Compulsory labor force participation did not emancipate women in Eastern Europe; as in the West, women had double burdens and the "second shift." Women felt themselves to be equal to men, however, because of their mutual oppression and lack of power. Child-care and maternity benefits were taken for granted; now they are rejected as part of a paternalistic package or as far too costly. Although women held many more political positions under socialism than at present, "tokenism" without power is now repudiated. Although the loss of reproductive rights met with substantial organized opposition, reversing these policies has proved difficult.

For feminists the critical question is What has been unleashed here and who is responsible? There are many potential villains: an aggressively political church; nationalist rantings; liberal, antistate policies that aim to dismantle the public sector. Each is partly to blame for the sudden increase in gender inequality, but as Einhorn acknowledges, too few women have fought the trends. Postsocialist media images verge on the utopian. Western television celebrates both consumerism and sexuality, in a manner that seems coarse and even vulgar when compared to the aesthetics (and ascetics) of socialism. Goods are equated with the good life; amazement and envy for the apparent universality of washing machines eclipses the question of who actually does the wash. Einhorn is skeptical that women will retreat to domestic roles or be convinced to worship at the altar of nationhood and bear larger families; at the same time, the drift is ominous. As in the West, women are more likely to be unemployed, and when employed they still tend to work for the state.

East-West confrontations have a way of challenging the conceptions of Western feminists as well. The antipathy to feminism is, at least in part, an aversion to individualism. An example will suffice. At the end of one of the first women's studies courses ever offered in Poland, the instructor was approached by an articulate young woman who solemnly announced that she could not accept the goals of the feminist movement, "because the moral and ethical foundations are flawed." 'What do you mean?' asked the astonished instructor. 'American feminism is based on seeking pleasure and choice,' responded the student quickly. 'And what should the ethical foundation be?' asked the instructor, taken aback. 'Self-sacrifice and an obligation to help others.'" Sorting out such moral and philosophical issues is important for politics and for feminism. This book is a good beginning for this task.

*The New American Workplace: Transforming Work Systems in the United States.* By Eileen Appelbaum and Rosemary Batt. Ithaca, N.Y.: ILR Press. 1994. Pp. ix + 287. \$45.00 (cloth); \$18.95 (paper).

Vicki Smith

*University of California, Davis*

This volume is the latest in a fine set of ILR Press publications that analyze cutting-edge changes facing the U.S. workforce. *The New American Workplace* reviews and synthesizes a wide range of quantitative and qualitative studies. The authors ask what configuration of organizational arrangements, public policies, and industrial relations might enable U.S. firms to move beyond the traditional mass-production model, facilitate competitiveness, and sustain a skilled and well-paid workforce. Comparative analysis, for Appelbaum and Batt, helps explain successful outcomes, so they draw extensively on data about Japanese and European systems (such as lean production, sociotechnical approaches, diversified quality production, and flexible specialization) to frame the discussion of U.S. alternatives to organizing mass production.

The authors optimistically argue that these non-American production strategies support high wages and high skills and depend heavily on unionized labor's cooperation and participation, thus increasing the bargaining power of the latter. But the bad news for the United States—both for workers and for everyone hoping for robust industrial development in an international, postindustrial economy—is that despite widespread innovation, U.S. experiments with workplace change do not advance a sustainable, systemic transformation that would achieve the advantages of the Japanese, Germans, Swedes, and Italians. The U.S. approach has been piecemeal, in Appelbaum and Batt's view, in that companies pick and choose from a "menu" of existing approaches. It is an eclectic approach that both exploits and favors low-skill, low-wage outcomes for labor.

Furthermore, the comparative data highlight the regrettable absence of a cohesive "institutional framework" for maximizing new workplace innovations in the United States. Appelbaum and Batt argue that because it lacks government policies that continually train the workforce and buffer workers from technological and organizational upheaval or that nurture research and development and interfirm coordination and because corporate profit strategies reward short-term speculative behavior and punish "patient" capitalists, American industry will be consigned to a haphazard and ad hoc industrial framework with deleterious consequences for its ability to compete.

The bulk of the book is devoted to mapping out the extent and character of innovations recently adopted by U.S. firms. Covering unionized and nonunionized settings, manufacturing and service industries, Appelbaum and Batt amply demonstrate that the current corporate emphasis on participation, flexibility, and restructuring goes far beyond rhetoric. They analyze in depth recent surveys and qualitative case studies, ac-

cessibly presenting the data about progressive workplace innovations in two long appendixes. These data suggest that, while the breadth of U.S. innovations is impressive, the depth of change is questionable. Firms may be counted as innovative if they have only one employee-involvement practice, and it is difficult to view this as evidence of deep-rooted transformation in, for example, a large, multidivisional company.

In addition to showing the range of new practices, the authors insightfully discuss current dilemmas in interpreting data about them, a problem experienced by any researcher who relies on the accounts of managers or on secondary reports found in the press. Definitional and methodological problems are substantial, such as the fact that managers in different companies may refer to very different arrangements with common terms like "employee involvement programs" or "self-managed work teams." Conversely, managers across a handful of corporate settings may lack a common vocabulary for describing what is essentially the same production arrangement.

Surveys often fail to tap the dynamics of change and involvement, while even in-depth interviews may distort accounts of the adoption and failure of programs because managers may try to portray new policies, and their role in promoting them, in the most positive light possible. (The book inadvertently presents a strong case for conducting in-depth field research to disentangle the multiple and complex dimensions of workplace transformation.)

It would be difficult to find more information about the depth and breadth of workplace change today than in this well-organized and comprehensive book. Yet reading it reinforces one's skepticism about the direction that our economic structure and the status of U.S. workers are taking. Appelbaum and Batt point out that despite great enthusiasm for long-term workplace reform—enthusiasm inspired by real problems in maintaining competitiveness—many employers, in the current economic environment, will gladly forsake reform to satisfy the short-term interests of investors and shareholders. This sets substantial limits on progressive corporate transformation in the United States. As long as the same corporate managers that adopt cutting-edge participative systems also lay off hundreds and thousands of their employees with little warning, they erode the trust and cooperation necessary for building long-lasting structural change in our production systems.

*The Change of a Lifetime: Employment Patterns among Japan's Managerial Elite.* By John C. Beck and Martha N. Beck. Honolulu: University of Hawaii Press, 1994. Pp. vii + 286. \$39.95 (cloth); \$18.95 (paper).

Hiroshi Ishida  
Columbia University

*The Change of a Lifetime* examines changes in the pattern of lifelong employment among Japan's corporate elites. The book focuses on a small

group of male corporate elites—graduates of “elite” colleges who are in career-track positions in “elite” companies—but it often reports data on the college-educated population as a whole and on trends throughout the entire labor market. Using documentary evidence, interviews, and a questionnaire survey, the authors arrive at two bold conclusions: first, that significant changes indicating a move away from lifetime employment have taken place in Japan since the late 1980s and, second, that the higher rate of changing jobs among elite corporate workers in the late 1980s is likely to remain stable “throughout the foreseeable future” (p. 249).

The book begins with the historical background of the lifetime-employment system (chap. 2) and its postwar development (chap. 3). These chapters, which rely heavily on a few selected sources (such as Thomas Rohlen's *For Harmony and Strength* [Berkeley: University of California Press, 1974] and Ezra Vogel's *Japan's New Middle Class* [Berkeley: University of California Press, 1963]), point out the importance of the traditional value of loyalty to the group in explaining the emergence and development of a “uniquely Japanese” system of lifetime employment. Because this system is firmly grounded in the Japanese tradition, according to the authors, it maintained integrity in the postwar period and contributed greatly to rapid economic growth.

The remaining chapters of this book describe how this system began to change in the late 1980s. The main argument of the book is that even though the lifetime-employment system remained firmly established in Japan in the late 1980s, both employers and employees became much more tolerant of the idea of job changing and a clear trend of increasing mobility was witnessed. Shifts in attitudes toward changing jobs are evident in many opinion surveys. More people were thinking about quitting their present job and willing to listen to offers from a headhunter in the late 1980s than in the past. Statistics related to changes in the actual number of job movers among elite employees are harder to find. Nonetheless, the authors rely on two kinds of statistics to support their claim. First, national-level statistics reported by the Ministry of Labor show that Japanese firms employed more male college graduates who already had work experience than those who had no work experience in 1988 and 1989, whereas the situation was reversed in the early 1980s. Second, the authors' own survey of 157 college graduates in the Tokyo area in 1988 shows that, among the 30 respondents who had changed jobs, the majority of them did so between 1986 and 1988. However, contrary to what the authors claim, the statistical evidence does not show conclusive trends. The national level figures include college graduates of elite and nonelite schools regardless of industry and firm size, and a sample of 30 job changers is not sufficient to derive a bold prediction about the Japanese lifetime employment system in the 1990s.

The most entertaining part of the book is perhaps the discussions of case studies of individuals who experienced job hopping and of firms that extensively recruited midcareer college graduates. These accounts

provide vivid descriptions of the psychological and social pressures that many job hoppers faced as well as the organizational struggles of the firms that were forced to accommodate the new workforce into the system of lifetime employment.

The book examines four macrolevel factors that contributed to the increased mobility of the Japanese workforce: the aging of the population, the slow pace of economic growth, the need for diversification, and the internationalization of the market. Given these macrolevel factors, the authors point out three kinds of organizations where midcareer recruitment proved to be successful in Japan: (1) foreign companies that aggressively recruited Japanese managers and engineers, (2) Japanese high technology firms and firms attempting to diversify that required engineers and technicians with specialized training and knowledge, and (3) large Japanese companies in the service sector (such as banking and security), where the demand for employees with international experience increased markedly. These analyses, which specify the context within which employment relations underwent significant changes in Japan, are the most informative part of this book.

It is unfortunate that the data presented in the book end around 1989. It is an open question whether the changes observed in the late 1980s continued into the early 1990s, when the Japanese economy experienced a major recession. Another limitation concerns female employees. Although Japanese corporate elites are predominantly men, female college graduates who are in managerial track positions are completely ignored in the book. The authors might have found more evidence for their claim of changes in Japanese employment relations if they had included female corporate elites since the Equal Employment Opportunity Law was enacted in the late 1980s. The book, nonetheless, reveals that the notion of the homogeneity and stability of the Japanese male corporate world is in need of heavy qualification.

*Union of Parts: Labor Politics in Postwar Germany.* By Kathleen A. Thelen. Ithaca, N.Y.: Cornell University Press, 1993. Pp. xii + 262. \$35.00.

Colin Crouch  
*University of Oxford*

Kathleen Thelen has given us, for the first time in the English language, a thorough account of how the German industrial relations system really works. Over the years it has been possible for this complex system to be portrayed in a number of more or less mutually inconsistent ways. The problem lies in its well-known dual nature: centralized collective bargaining alongside company-level work councils. Such a model can be perceived—as indeed German trade unions perceived it for most of the first two decades of its existence—as one where unions pursuing normal

antagonistic collective bargaining are constantly undermined by a collaborative and nonunion form of consultation through works councils. Alternatively, by skipping over the differences between unions and works councils, it can be depicted as a model of legally guaranteed union workplace strength.

What really seems to happen, at least in those industries where unions have learned how to use it, is a subtle and complex balance between the different limbs of the dual system. And to understand fully what is meant by balance here it is necessary to follow Thelen's reasoning and reject the automatic assumption that intraorganizational relationships between center and periphery are necessarily zero sum. Works councils *do* constrain unions: They take on a number of tasks that otherwise might be performed by unions, reducing the latter's scope, and they are bound to a peace obligation with the employer. On the other hand, unions have very often achieved *de facto* penetration of the works councils by having their candidates elected to them, though this in no way guarantees that those elected will then automatically follow central union preferences in their work in the firm.

Unions have to work hard to secure the cooperation of works councils; they therefore incorporate works councillors in their bargaining commissions and decision-making bodies. This incorporation integrates them in the union, binding the union to the perspective of works councils.

The whole process is an example of the subtle *Verflechtung* (interweaving) that is at the heart of so many modern German institutions and is so different from the stereotypes of that country. At the same time, a good deal of the process has been largely accidental. Although Germany gives the impression of having a kind of designer industrial relations, put in place as a deliberate new conception in the years after World War II, no one ever envisaged the dual system working quite the way it does.

Although she does not say so explicitly, Thelen is also contributing usefully to collective action theory, providing an example of the limitations of simple central/decentral dichotomies. There *is* a strong coordinating capacity at the heart of German trade unions and, therefore, collective bargaining. But it is a center that is required to take serious note of the periphery, in a process from which both levels (and the several intermediate ones) gain. And if there is tension between them, it is not necessarily that routinely envisaged in social theory of a center that is capable of seeing national priorities and therefore becomes detached from members, while local representatives remain close to the latter's limited, nonstrategic perspectives. The company consciousness imposed on local representatives under a works council system requires them to be something more than defensively reactive. Thelen expands very interestingly on this last point by contrasting German and U.S. industrial relations. Both countries' systems would rank high in any comparative assessment of degree of legalism, but the American one, she argues, is contractually based, while the German is constitutionally anchored. One consequence

of this system is that German unions do not have to protect their established rights by defending all past victories but can risk bargaining for innovation.

She demonstrates this last characterization of the German system by showing how German unions were able to bargain their way through the major restructuring crises of the 1970s and 1980s. Confident nationally in the constitutional base of their rights, responsive locally to companies' needs, they were able to adopt modernization strategies. The complex interrelationship of center and periphery enabled them to operate with remarkable flexibility while still developing general strategy. This is demonstrated through case studies of three important industries: steel, automobiles, and consumer electronics. These are very diverse, and the story is by no means the same; this account is therefore as nuanced at the detailed level as it is clear in the strategic political implications.

Overall the book presents a favorable view of the achievements of German organized labor, demonstrating how it has been able to behave in ways usually only associated with small countries, where a strategic level of action is that much easier to achieve, while giving full attention to the ways in which it differs from them. But it is not uncritical. Among the three industries considered, the story in consumer electronics is less successful than in the other two. Nationally, Thelen shows how Germany was unable to sustain levels of employment as high as in several other advanced economies. It is notable that, using the earlier work of Fritz Scharpf ("Economic and Institutional Constraints of Full-Employment Strategies," pp. 257-90 in *Order and Conflict in Contemporary Capitalism*, edited by John H. Goldthorpe [Oxford: Clarendon, 1984]), she is able to attribute this not to restrictions imposed by German labor practices but to the behavior of the Bundesbank.

*The Scar of Race.* By Paul M. Sniderman and Thomas Piazza. Cambridge, Mass.: Harvard University Press, Belknap Press, 1993. Pp. ix+212. \$18.95.

David O. Sears

*University of California, Los Angeles*

This book, on the role of race in the political attitudes of the white mass public, is something of a compromise between an academic research report and a trade book critiquing contemporary social scientific writings from a neoconservative viewpoint. Its main message is that the civil rights agenda, fully harmonious with "the American Creed" in the days of Gunnar Myrdal, now conflicts with it, thereby generating much non-racist white opposition (especially to affirmative action).

The book takes on large questions. It reports a study that was innovative in a number of respects. The book takes a commendably pluralistic,



often imaginative, approach to empirical testing, employing both correlational and experimental techniques rather than relying solely on the often overused multiple regression of survey data.

However, the decision to gear the book for a lay audience may have mixed consequences for its reception. The book will no doubt receive more than the usual attention from public intellectuals, particularly those ideologically in tune with its main message. As a piece of scholarship in an intensely researched field, however, the book is dissatisfyingly non-cumulative, rarely placing the authors' new results in the context of the extensive empirical literature on exactly the same problems (other research on racial attitudes is cited fewer than a dozen times in toto). Also, its style of presentation may be too nonspecific and methodologically simplistic to allow scholars in the field to properly assess, much less accept, the authors' claims that the book introduces "a radically new way of conducting studies of public opinion" (p. 11), addressing "questions that can, for the first time, be answered" (p. 67), with findings that uncover "a sea change in the temper of racial politics" (p. 23).

The book makes three specific claims. First, racism has a much smaller role in mass white politics than many others think it has. There is no "new racism" or "covert racism," just a familiar general ethnocentrism (derived from authoritarianism) that manifests itself variously in old-fashioned racial stereotypes (especially perceptions of blacks' lack of work ethic) and anti-Semitism. Second, racial prejudice has only modest influence over attitudes toward racial policies (which instead are assessed on their own merits, issue by issue). Third, not being rooted in an unvarying racial prejudice, these policy attitudes are neither highly consistent across issues nor unusually resistant to change; rather, support varies across issues and responds to good arguments, allowing political coalitions to be made and remade as policy proposals change.

Three general solutions are proposed for the long-standing racial problems of our society. Blacks need to become more responsible and hard-working to counter whites' reality-based stereotypes; whites need to be better educated, generating more tolerance as they tie their policy attitudes more closely to their moral principles; and elite liberals need to generate more sensible policy proposals.

Much of the book is framed as a long series of refutations of others' claims. These adversaries are generally described, without citation, only in such shadowy terms as "contemporary commentators on race," "the conventional wisdom," "the new racism researchers," "the consensus view," "the common suggestion," or "recent claims." Since they are not usually identified, it is usually unclear whether these adversaries are social scientists, legal theorists, legislators, political activists, or op-ed writers.

The lack of citation is more than just a matter of academic nicety, because the claims attacked by the authors often resemble straw men more than they do widely held theories: for example, "the response of

whites to specific issues of race, it is said, is so massively consistent that plainly the particulars of specific issues and policies are not, in themselves, important" (p. 18); "White Americans respond . . . to the contemporary array of racial issues *en bloc*, either opposing or favoring them across-the-board" (p. 25); "Opposition to affirmative action is said to be, in and of itself, racism" (p. 97); "The new racism researchers' thrust has been to assert that ideology reduces to racism" (p. 120); "It is arresting to hear that individualism has become, in our time, the source of racism" (p. 173-74); and "It is a mistake, a deep and disfiguring mistake, to believe that racial politics does not take part in regular politics" (p. 176). No well-known social scientist who works on racial attitudes holds these views, so far as I know.

Do the authors make a good case for their own claims? First, do racial policy attitudes on different issues cluster in various separate issue agendas, no longer driven by a simple underlying racial prejudice generating consistent antiblack (or problack) viewpoints? The authors do distinguish three separate "agendas" but present no direct test of cross-issue inconsistency, such as a factor or cluster analysis, nor any evidence that such inconsistency has increased over time.

Does racial prejudice no longer have much influence over racial policy preferences? Again, the authors present no comparisons over time. Currently, however, opposition to the "social welfare" agenda (e.g., the call for more government spending on blacks) is linked to whites' stereotypes of blacks' lack of work ethic and to both prejudice and general conservatism (p. 117). Opposition to the "equal treatment" agenda (e.g., fair housing) stems from prejudice, not ideology (pp. 123, 126). The authors do not assess the origins of opposition to the "race conscious" agenda, because they say that for busing and affirmative action "there is scarcely any support for either among whites" (p. 130; notwithstanding the data showing that 25% of whites support each policy and 50% support minority set-asides; see pp. 131, 133). These findings resemble those from many other studies: strong prejudice effects for all racial issues, with general political racial ideology or nonracial traditional values making an independent contribution on more contemporary issues.

Will whites change their racial policy preferences, given good arguments? Some experiments are presented that report attitude change following brief counterarguments. The commendable originality of this approach is offset to some degree by selective data reporting and by an absence of significance tests, tests of the durability of change, or reference to the extensive literature showing impressive stability of racial attitudes over time.

Are racial stereotypes openly expressed rather than covert? The stereotype of blacks' poor work ethic is still common and openly expressed, as both Kluegel ("Trends in Whites' Explanations of the Black-White Gap in Socioeconomic Status, 1977-1989," *American Sociological Review* 55 [1990]: 512-25) and Bobo and Kluegel ("Opposition to Race-Targeting:

Self-Interest, Stratification Ideology, or Racial Attitudes?" *American Sociological Review* 58 [1993]: 443–64) reported several years ago from GSS data (in work not cited by Sniderman and Piazza).

Do racial stereotypes stem from generalized ethnocentrism, as in the theory of the authoritarian personality, rather than specifically from anti-black affect? The authors do find racial stereotypes correlated with anti-semitism, treated here as a surrogate for ethnocentrism. The authors argue that such stereotypes are also rooted in whites' real direct experiences with blacks: "It is part of the ordinary experience of many whites to encounter blacks who are not trying as hard as they could to overcome their problems" (p. 106). However, no data are presented on whites' personal experiences with blacks (or with any other group) or on the linkages of such experiences to whites' policy preferences. This observation itself risks being more stereotype based than data based.

In the end, no evidence is presented contrary to Converse's ("The Nature of Belief Systems in Mass Publics," pp. 206–61 in *Ideology and Discontent*, edited by D. E. Apter [New York: Free Press, 1964]) original demonstration that racial attitudes are more stable and consistent than other attitudes. And it is clear that racial prejudice is strongly linked to racial policy preferences. Many scholarly readers may be troubled by the concessions the authors made to trade book presentation. The book seeks to stimulate controversy, which is almost always healthy for scholarship. But effective challenges to others' views in intensely researched areas are more likely to come from better documented and more convincing evidence.

*Sharing the Journey: Support Groups and America's New Quest for Community.* By Robert Wuthnow. New York: Free Press, 1994. Pp. xiii + 463. \$22.95.

David M. Hummon  
*Holy Cross College*

In *Sharing the Journey*, Robert Wuthnow invites a diverse audience to explore the character and significance of the small-group movement in contemporary American society. Speaking to movement participants and clergy as well as academics, he offers a highly readable analysis of the contributions of support groups to individual well-being, spirituality, and community life. Although sympathetic to the movement, Wuthnow writes not as an apologist but as a scholar, concerned with understanding the movement as it both enriches and limits human experience.

*Sharing the Journey* combines survey analysis with qualitative research to construct an exceptionally rich interpretation of small-group life. Using a representative sample of the U.S. public that includes more than 1,000 small-group members, Wuthnow provides the first national analysis of the demographics of group membership, of the varied types

of groups, of the motives people have for joining groups, and of group dynamics. Drawing on the work of 15 research associates, Wuthnow also incorporates ethnographic interviews and observations into his presentation. These materials enable him not only to illustrate survey findings with flesh-and-blood examples but also to explore questions of group process and cultural meaning.

Wuthnow offers compelling evidence that small groups have become a significant feature of the U.S. landscape. Currently, 40% of the U.S. public report that they participate in a small group "that meets regularly and provides support or caring for those who participate in it" (p. 395). Moreover, such participation is distributed across the diverse groups that compose U.S. society. Though women, older individuals, and better educated Americans are somewhat more likely to join groups, such demographic differences are modest, and other social cleavages that frequently divide Americans—race, income, region, community size—are even less consequential. Such inclusive participation is encouraged by a wide diversity of groups. Of contemporary support groups, Wuthnow estimates that 29% are adult Sunday school classes, 25% are Bible study groups of different forms, 12.5% are self-help groups (4 in 10 of these are related to addiction), and the remaining 33% are special interest groups (these are divided equally among groups devoted to book discussion, politics and current events, and sports and hobbies). Though members of these diverse groups report receiving many types of support from their groups, these groups excel in providing communal and emotional support, helping participants to overcome feelings of isolation, providing encouragement, teaching self-acceptance, and nourishing self-esteem.

Wuthnow attributes the growth of the small-group movement during the last three and a half decades to its cultivation of a sense of community and of the sacred. Although Wuthnow places this quest for community in the context of a breakdown of such traditional sources of communal ties as neighborhood, he argues that support groups are expressive of the very social forces that make them personally advantageous and socially adaptive. At once formally organized groups with leaders, goals, and schedules and at the same time small, voluntary, relatively fluid, non-demanding, and caring, support groups paradoxically provide a formally structured informality, a rationalization of intimacy. As such, they not only offer a sense of community to members but they also transform the very ways that Americans experience and think about community. Such a transformation, Wuthnow suggests, may well involve social and personal costs, such as a weakened sense of the nature of commitment and a narrowing of social horizons.

Wuthnow emphasizes that the growth of the support-group movement must also be understood as an expression of people's search for the sacred. He argues that this function of small groups is not indicative of secularization but of the very strength of contemporary religious life. The support-group movement has been actively supported by church organizations, which have provided facilities, leaders, and a common

culture of caring. Many Americans join support groups explicitly to enrich their faith or grow spiritually, and members report that they feel spiritually nurtured by their groups. Although such enrichment is cultivated by group opportunities to discuss religious texts or pray, Wuthnow documents that such growth is particularly encouraged by providing a context in which personal stories are told and imbued with spiritual meaning. Here again, however, Wuthnow suggests that the small-group movement's engagement with spirituality may be transforming the very ways that Americans conceive of the sacred. For example, the intimacy, caring, and self-directed nature of support-group culture may be domesticating conceptions of God, cultivating a God of love rather than obedience, of personal needs rather than social justice, of immediate, pragmatic ends rather than transcendence and worship.

Wuthnow's interpretation of the support-group movement in terms of religious and communal dimensions of U.S. life is both complex and provocative. Though support groups clearly serve the needs of the individualistic self of contemporary U.S. culture, Wuthnow demonstrates that such groups also encourage members to become more active in their community through volunteer work. Such findings suggest that support groups may sometimes function as mediating institutions, turning Americans outward toward the larger society as well as inward toward the self. Such arguments, theoretically informed and empirically grounded, make *Sharing the Journey* a compelling social analysis of a significant facet of U.S. life.

*Goodness Personified: The Emergence of Gifted Children.* By Leslie Margolin. New York: Aldine de Gruyter, 1994. Pp. xxiii + 181. \$37.95 (cloth); \$18.95 (paper).

Spencer E. Cahill  
*Skidmore College*

Enrollments in social problems and deviance classes attest to the allure of the socially forbidden and troubling. But Leslie Margolin's fascinating and informative book should turn some heads, for it perceptibly reveals the scholarly charms and rewards of focusing attention on the socially valued.

Margolin takes the reader on an intriguing historical journey. The book begins by discussing the eugenic origins of the gifted-child movement, follows the rhetorical twists and turns of the movement's scholarly allies, and passes through an all too brief ethnographic chapter about a counseling center for gifted children and their families. There are numerous stops along the way to theoretically digest the empirically lush panorama.

As Margolin documents, the founders of the gifted-child movement reportedly discovered not only cognitively but also physically, morally,

and spiritually gifted youngsters. These founders were not troubled that these little *ubermensch* almost invariably came from affluent, respectable, and lily-white families. It made perfect scientific and common sense to them. Exceptional goodness was an inheritance rather than an accomplishment. Margolin then shows how gifted-child scholars later renounced their forerunners' unfashionable class and racial bias without renouncing the core belief that goodness is endowed. Admirably self-critical and gracious in the face of criticism, they rhetorically ducked responsibility for what they and their forerunners had wrought. For example, they shifted the blame for the underrepresentation of African-Americans among gifted children from faulty genes to dulling environments, safely avoiding taking any blame themselves. Yet, despite their self-serving rhetoric, Margolin does not question the sincerity or motives of gifted-child crusaders. Rather, he convincingly argues that they, like everyone else, are oblivious to their own reality-constructing work and its consequences.

However, Margolin makes it abundantly clear that gifted-child scholarship has been and remains passionately partisan. Its purpose is to promote special and segregated education for gifted children as variously defined and identified. Over the years, gifted-child scholars have argued that the exceptional talents of gifted children are squandered in classrooms with ordinary students and teachers, where they are woefully misunderstood. They also question parents' ability to keep up with their prodigious offspring. Margolin is justifiably impressed by these scholars' rhetorical success and genius. They created a victim without weakness—the gifted child—and a type of villain who is fully redeemable—the ordinary and uninformed peer, teacher, or parent.

As Margolin documents, no one is more persuaded by the rhetoric of the gifted-child crusaders than its authors. They are oblivious to the possibility that exceptional educational opportunities might produce gifted children. They are obviously silent about the possibility that ordinary children might benefit from those same educational opportunities. And they are oblivious to the possibility that their success in securing public funds for gifted children's special education might be at the expense of their more ordinary peers. Like their predecessors, contemporary crusaders for gifted children and the professional experts who nurture the goodness of its young personifications are convinced that goodness is endowed. As Margolin illustrates with transcribed excerpts of counseling sessions recorded at a center for gifted children, once a child is designated gifted, those professionals view almost everything he or she does as symptomatic of his or her gifts. Through Margolin's social constructionist lens, the counselors' tortured interpretations of their gifted clients' words and deeds are almost comical. Yet the counselors are deadly serious, and that sobering thought dilutes their unintended humor.

Margolin's social constructionist parable has many theoretical morals. I will only mention two here. First, Margolin's astute analysis of how the socially constructed category of gifted children diminishes their peers

suggests that we can learn much about the socially bad, ugly, and just plain ordinary from studying the social construction of goodness. As Margolin illustrates, images of goodness create, by implicit contrast, images of evil and mediocrity. Second, Margolin's critical examination of educational programs for gifted children confirms Foucault's contention that the exercise of power has become less punitive and more seductive and, consequently, more effective. Those programs entice conformity with the lure of enhanced status and the promise of future success. As Margolin argues, it is a temptation and power that few children or parents can resist.

I have not done justice in this brief review to the empirical bounty, theoretical insights, and scholarly pleasures of this thin volume. Though Margolin repeatedly interrupts the reader's pleasure by dropping names and by relying too heavily on these persons' arcane expressions, this is a minor irritation. Readers who politely ignore it will be richly rewarded. This book is a treat and deserves a wide audience among students of social life, whatever their substantive interests. There are lessons for all of us in its pages.

*Leaving Home before Marriage: Ethnicity, Familism, and Generational Relationships.* By Frances K. Goldscheider and Calvin Goldscheider. Madison: University of Wisconsin Press, 1993. Pp. xvi + 242. \$27.50.

William G. Axinn  
*Pennsylvania State University*

The rapidly rising incidence of nonfamily living over the past four decades is one of the most dramatic changes to sweep American society in recent times. Although the extent of this change has been documented, relatively little is known about the factors that have produced this dramatic change. Furthermore, evidence suggests that nonfamily living has important consequences for many areas of social life, including both family formation and gender relations, so that understanding the causes of nonfamily living is a crucial step toward explanations of many recent social changes. In their book *Leaving Home before Marriage*, Goldscheider and Goldscheider turn to individual-level data (the High School and Beyond surveys) to explore the factors that are associated with nonfamily living in early adulthood. The Goldscheiders ground their exploration in theories of ethnicity, familism, and intergenerational relationships, so that the individual-level results provide important clues about these aggregate-level social changes. The book provides the most comprehensive examination of the causes of variations in nonfamily living to date, and as such it is important reading for a broad range of social scientists.

The chapters of *Leaving Home before Marriage* begin with a description of premarital residential independence (PRI) and the relationship be-

tween expectations for this independence and actual behavior. Then Goldscheider and Goldscheider build a systematic causal model of factors affecting PRI. They begin with factors that are exogenous to the lives of young adults, namely their racial and ethnic heritage, and build toward more proximate causes of young people's behavior, including religion, religiosity, and attitudes about relationships to parents, gender relations, and marriage. The later chapters address intergenerational effects on PRI by examining the effects of family structure, parental resources, intergenerational resource transfers, and parental expectations. The book provides parallel analyses of expectations and behavior that continually emphasize the factors that influence both or one and not the other. Results build toward the general theme that family-oriented values, the family context, and other family members (parents) are central determinants of individuals' expectations of and experiences with nonfamily living.

The highlights of the book are a number of findings that reach beyond nonfamily living per se to provide insights into the causes of behavior more generally. The book joins previous studies by demonstrating that young people's expectations are important determinants of their subsequent behavior, even when controlling for factors that predict both expectations and behavior. The analyses also show that parental expectations are an important determinant of young people's behavior, even when the young people's own expectations are taken into account. This work also adds to the literature indicating that attitudes about a range of family relationships are important influences on young people's lives. Thus, the results demonstrate the important role family context plays in shaping the lives of young people, including subjective aspects of family context.

The most conspicuous omission from this book is the word "cohabitation." The same decades that have seen the dramatic rise in nonfamily living have seen an equally dramatic rise in nonmarital cohabitation. Many of the factors that the authors show increase the likelihood of PRI also increase the likelihood of premarital cohabitation. The authors' definition of PRI appears to include cohabitators, so it is difficult to tell whether results are primarily produced by effects on cohabitation or effects on other forms of residential independence (living alone or with housemates). The addition of cohabitation to both the conceptual framework guiding the analyses and the analyses themselves would have clarified and strengthened the book's contribution. On the other hand, recent work on living arrangements suggests substantial fractions of those who experience a premarital cohabitation experience residential independence first. Thus, analyses of cohabitation would also be strengthened by taking other forms of PRI directly into account.

While providing as comprehensive a picture as possible, given the limitation to one data set, *Leaving Home before Marriage* also provides an atlas filled with tantalizing directions for new research. Additional measures of family values and family relationships are likely to enhance



our understanding of the processes leading to residential independence. In fact, a better understanding of familism may improve our insights into many aspects of family behavior. The authors' observation that living arrangements are at least as much household-level decisions as individual-level decisions should spur researchers to develop better theories of family and household influences, along with collecting more comprehensive data from other household members. The authors' prediction that levels of nonfamily living are likely to rise should also motivate more aggressive research on the consequences of nonfamily living. Of course, research on these consequences needs to distinguish between nonmarital cohabitation and other forms of nonfamily living.

Undoubtedly this book will be the starting point for these and other important avenues of future research. It provides a comprehensive examination of the determinants of nonfamily living, insights into the causes of behavior more generally, and an outstanding foundation for further research. Therefore I recommend it highly.

*Out of the Garden: Toys and Children's Culture in the Age of TV Marketing.* By Stephen Kline. London: Verso, 1993. Pp. x+406. \$34.95.

Gary Alan Fine  
*University of Georgia*

Some titles say it all. Readers of this journal will already have guessed what the theme of Stephen Kline's engaging, if unsurprising, volume is. The snakes in Kline's Eden are marketing managers. Kline (p. 23) avers: "The shadow that marketing casts upon our most cherished means of cultural expression is now too ominous to be viewed naively." Those who cherish their naïveté have only themselves to blame. To suggest that *Out of the Garden* lacks the element of surprising revelation, of course, does not suggest that the work is without interest, nor does it demonstrate that Kline's thesis is inaccurate. That Chicken Little was overly worried does not mean that there is never a cause for alarm.

For Kline the fundamental threat from marketing involves its invasion and alteration of children's culture, not the economic vulnerability of youth. As many others have done before him, Kline frets that our contemporary culture of consumption, particularly the animated Saturday morning visual ghetto, has corrupted a purer, precapitalist children's culture. This precapitalist culture did not, as Kline notes, disappear all at once but has been vitiated by a gradual process in which children's innocence has been corrupted by encroachments of external, materialist, adult forces. One of the strengths of Kline's analysis is the space devoted to the marketing of play and children's literature prior to the advent of television culture and spinoff toys.

The decision of whether to embrace Kline's vision—with its attendant values—is perhaps more a matter of preference than of evidence. Surely

one cannot deny that children's culture has been altered over the centuries, decades, and days, but the specification and evaluation of this change is a greater challenge.

The frequently presented argument that children's culture used to be cut off from the culture of adults is difficult to embrace, because children have simultaneously been incorporated *and* expelled from the domains of adult life. Children have been incorporated into both adult media and adult-dominated institutions (schools, day care, youth sports, camps); yet, simultaneously, and without grace, they have become increasingly removed from the community of adults. Children are colonized in their own specialized institutions, guarded by adults paid for that task. The multiage community of the streets has been lost except in some quadrants of the inner city.

Children develop, as they always have, a relationship to the culture of the adult world, but the relationship is more instrumentally grounded than expressively linked. Children do not find themselves in the world of adults, so much as they now find adults crouching at the edges of their world, controlling, protecting, and tempting. Kline does not argue that children should be brought into the world of adults, nor does he wish that children be left alone to make their own market choices; rather he embraces the legitimacy of adult regulation of children's culture—switching one set of adult controllers for another. Such regulation need not be a bad thing, but it is most explicitly an ideological choice—a choice that is not grounded on demonstrable harm to children but on demonstrable affronts to adult beliefs of how childhood should be organized.

At its best, *Out of the Garden* explains the diversity of the themes of children's culture. Kline reminds us that the representations of popular culture aimed at the young permit children to embrace a diverse range of identities; he properly is suspicious of those who suggest that all children's culture is violent or materialistic. Yet, he retreats into the romantic complaint that we have failed “to make the marketplace a positive cultural force in contemporary society” (p. 350), casually assuming the definition of what such a positive force might be.

If the world of children's culture is a true market, then it should respond to the demands of its customers. The fact that it is not satisfactory for many adults suggests either that it is a free market and the problem is with the consumers, or that it is not and the problem is with those elements that prevent it from being a true free market (where the most desired products would survive), or that it is some combination of the two.

To the extent that the problem is with the child consumers, then adults must be involved to “inspire children with high ideals or positive images of the personality, provide stories which help them adjust to life's tribulations or promote play activities that are most help to their maturation” (p. 350). Kline seems to say that adults must provide children with what they need but not what they would freely choose. The media, as improper

guides, have structured the identity and imagination of children, blunting creativity. Yet, this bracing notion cuts against the image of the volume—from the title on—that children have been led “out of the garden” by amoral marketers. The question becomes not Will children be led? but Who will be the Pied Pipers?

If the problem is the market, then perhaps we should permit a greater flowering of competition in which successful products will be available to be chosen by the consumers themselves. Do we have enough faith in children to let them make their own choices?

Most subversively, some might suggest that perhaps the world we live in is, despite our anxieties, the best of all possible worlds: a romanticism of media-saturated kids. The growth of children's culture, from this perspective, does not demand our concern.

*Careers and Creativity: Social Forces in the Arts.* By Harrison C. White. Boulder, Colo.: Westview, 1993. Pp. xviii + 219. \$55.00 (cloth); \$16.95 (paper).

Diana Crane  
*University of Pennsylvania*

*Careers and Creativity* is intended to provide an introduction to the sociology of art for undergraduates. The book focuses on interactions between the artist and his (seldom her) social and cultural environment and relationships between social identities and the arts.

The book proceeds by examining a variety of cases that exemplify White's theses in various ways. His first example is that of the Pre-Raphaelites, a group of 19th-century English artists and craftsmen. White argues that the Pre-Raphaelites' conception of their artistic identity shaped the public's perception of its own class identity. In chapter 3, he presents his perspective on the relationship between personal identity, narrative, style, and artistic careers. To paraphrase a statement later in the book, the arts are a celebration of identities and an accounting of relations between identities (p. 121). I suspect that undergraduates will find this chapter particularly dense.

In chapter 4, White deals with major shifts in style, mainly in the arts but also briefly in popular culture, arguing that stylistic changes require changes in the social organization of art worlds. Here he draws from his classic book on 19th-century French artists (Harrison C. White and Cynthia A. White, *Canvases and Careers* [New York: Wiley, 1965]) and case studies by other authors (e.g., M. Baxandall, *The Limewood Sculptures of Renaissance Germany* [New Haven, Conn.: Yale University Press, 1980.]). The following chapter presents notions of agency and the ways in which minorities and ethnic groups, for example, use the arts to construct social boundaries.

Chapter 6 analyzes the problems of a particular performing art, the

theatre. Curiously, White attempts to map the institutional system of the U.S. theatre and its current dilemmas without discussing DiMaggio and Stenberg's extensive study of the resident theatre ("Conformity and Diversity in American Resident Theatres," pp. 116-39 in *Art, Ideology, and Politics*, edited by J. Balfe and M. J. Wyszomirski [New York: Praeger, 1985]).

Chapter 7 uses case studies of artists who are former students and personal friends of the author to examine the problems of artistic careers. Simpson's systematic study of artists' careers in SoHo (*SoHo: The Artist in the City* [Chicago: University of Chicago Press, 1981]) is cited but not discussed. The fact that opportunities for artistic careers have generally been much less available to women than to men is ignored. In the concluding chapter, White laments the lack of support for the arts in the United States but makes no reference to the extensive literature on arts policy that documents and elaborates this theme.

The arts and artists are presented as unambiguously good for society and for the individual (e.g., "No ordinary book like this one can adequately convey to you the magic of art's levels of reality" [p. 191]). There is no discussion of the fact that, because the arts are supported by social elites, they typically support the interests of social elites. The ambiguous relationship between avant-garde artists and social elites (e.g., government and business), which has been the subject of numerous articles and books, has been deliberately omitted (see p. xv).

*Careers and Creativity* provides a highly selective view of the field using a writing style that is variously colloquial, patronizing, often abstruse, and punctuated with lengthy quotations from other authors. The book contains many important and interesting ideas, but too often they remain too abstract, in spite of copious examples, or insufficiently developed. The book might be useful as a text in an undergraduate course on the sociology of art in conjunction with other readings. Discussion of popular culture is too brief to justify its use in a course on popular culture. One nice touch at the end of each chapter is a list of suggestions for field studies students might do themselves, but the lack of summaries and conclusions at the end of chapters is a serious handicap in a theoretically oriented text aimed at an undergraduate audience.

At times, White appears to have set himself the task of creating a sociology of art, as if it did not already exist. Sociological studies that would have illustrated the points he is making are often ignored. The author invokes art-worlds and production-of-culture perspectives, but relevant literature from these approaches is often ignored. Throughout the book, White prefers to discuss humanists and the unpublished writings of his graduate and undergraduate students rather than social scientists; some relevant studies by the latter are cited in lists of recommended readings at the end of each chapter. This is an approach one might have taken 20 years ago, but, given the extent of the development of the sociology of art since then, it is a pity that an exciting and important field was not better served by this author.

*All Things to All People: The Catholic Church Confronts the AIDS Crisis.* By Mark R. Kowalewski. Albany: State University of New York Press, 1994. Pp. viii + 167. \$14.95.

Jim Wolf  
*University of Chicago*

Outsiders to the Roman Catholic Church are frequently puzzled by and often critical of the apparent contradictions within the church. The church hierarchy pronounces clear dictates regarding birth control and extramarital sex, while the laity apparently ignores them. The same priest who preaches that homosexual activity is a sin might be among the most active proponents of compassion toward gay people with AIDS. How should we understand these discrepancies?

In his book *All Things to All People: The Catholic Church Confronts the AIDS Crisis*, Mark Kowalewski offers evidence that these paradoxes are not just explainable but are the necessary means by which the church can effectively deal with the AIDS crisis and still remain true to its official teachings. The main dilemma of the church regarding AIDS is to fulfill its duty to minister to people with AIDS (particularly gay people) while standing fast on its policy regarding homosexual activity. Priests, because they provide both teaching and pastoral care to the laity, hold an important intermediate position. In their position as mediators between the policy-setting hierarchy and the faithful who are expected to follow those policies, priests play the necessary role of "good cop." This allows disenfranchised members of the church to come to a negotiated reconciliation with the church if such an end is necessary to assure their continued commitment to the church as a whole.

Kowalewski begins his work with a brief discussion of the church as a complex organization exemplifying a traditional authority structure. He discusses how priests maintain stability in the organization by making compromises as needed in their interactions with the laity. The laity view priests as holders of church authority bestowed upon them by the church hierarchy. He discusses the social and medical contexts that were current in the late 1980s, when most of the work was written.

Using an ethnographic approach, Kowalewski conducted 32 60–90 minute interviews with priests in the Los Angeles area in 1987–88. Most of the interviewees were known to the current director of the AIDS ministry for the archdiocese as being active in ministering to AIDS patients. Rather than start with specific hypotheses formulated in a particular theoretical framework, Kowalewski chose to use a "grounded theory" approach, allowing the theoretical construct to emerge from the data (Barney G. Glaser and Anselm Strauss, *The Discovery of Grounded Theory* [New York: Aldine, 1967]). The author also reviewed many published statements from the church hierarchy and included them in his analysis.

Kowalewski's review of official documents related to AIDS revealed that reactions by the U.S. bishops diverged from those of the Vatican.

Papal statements still asserted the air of *Roma locuta, causa finita* (Rome has spoken, the matter is settled). In other words, those who wish to be saved and welcomed back into the church must profess the evil of their ways and repent. American bishops were less demanding, focusing much of their teaching on the need to care for those with AIDS and not treat them with disdain.

When forced to educate about AIDS, most priests do not go beyond the official teaching of the church. While they will not publicly promote the use of condoms, they will denounce the treatment of people with AIDS as social lepers and promote more compassionate understanding of their situation. Preventive education was not a high priority to the priests interviewed.

Perhaps the most important activity performed by church members is ministering to people with AIDS. Priests perform various roles in this regard: companion, counselor (for the patient and the patient's family), spiritual guide, and minister. When challenged by the situation of helping a gay person with AIDS who requests reassurance, most priests interviewed resort to a negotiation strategy in which they search for some common ground between the patient and the church. The conclusions reached might never be professed from a pulpit, but they fulfill the priest's desire to provide care to a dying person in need of ministry. The result is that "while organizational administrators set official policy, such policy is then translated by practitioners into working knowledge so they can implement organizational goals and 'get the job done'" (p. 113). The author points out that while some may view the activities of priests as evidence of the decline of church authority, such activities are not considered evidence of a change in policy until those policies are changed by church officials.

Kowalewski's work provides a valuable look into two dimensions (public teaching and pastoral care) of how the church functions in the face of a moral crisis. His theoretical arguments meander somewhat, but this probably results from his grounded theory approach. In sum, this book will be of most use to those interested in gaining insight into how apparent contradictions within the church are actually necessary reactions to a crisis that threatens its internal stability.

*The Social Control of Religious Zeal: A Study of Organizational Contradictions.* By Jon Miller. New Brunswick, N.J.: Rutgers University Press, 1994. Pp. xvi + 238. \$48.00.

Adam B. Seligman  
*University of Colorado*

While there are many Weber scholars, there are few practitioners of Weberian sociology. Jon Miller is one of that dwindling number of sociologists who take Weber seriously, that is, who use Weberian methods and

concepts as working tools of their trade, as heuristic devices to explain the social reality they confront in their research. Setting out to explain the persistence—for close to two centuries—of the Pietistic Basel Mission to Ghana, Miller's analysis uses a Weberian problematique in the most illuminating and interesting of ways. It is therefore a book not of Weberian theory but of Weberian sociology as utilized by a fine craftsman.

Miller's analysis is oriented toward elucidating the interweaving of ideas and structures, to the way a certain set of beliefs—in authority, discipline, the establishment of a new "community of belief" and the role of Christianity in black Africa—influenced the mission's organizational structure and the consequences thereof. This analysis is carried out against a lucidly presented background of the mission's history (including similarities with and differences from other missions) and—what was of particular interest—a detailed intergenerational study of the mission's membership. The author's class analysis of both the mission leadership and the rank and file and also his joining of "ideal" with "material" interests in explaining what motivated servants, shoemakers, weavers, cabinetmakers, bakers, and farmers from the villages and towns of southern Germany (especially Wurttemberg) to emigrate as missionaries to sub-Saharan Africa is a model of what this type of analysis should be.

While not reducing the evangelical vocation to a crude set of material determinants, Miller's analysis provides a crucial dimension to our understanding, not only of the mission's place within the broader transformations of the 19th century, but of the organizational constraints and tensions of the mission itself. Miller's explanation of the missionaries in terms of the "search of many young Wurttembergers for a productive and moral role in a materially changing world" (p. 47) exemplifies well his sensitivity to the ideal interests of the social actors while not abstracting them for the very real material and social structural conditions within which they lived their lives and sought to maximize their goals.

Perhaps the heart of the book is the chapters devoted to authority, discipline, and the contradictions of organizational life. The author presents a careful study of how different modes of legitimation—charismatic, traditional, and legal-rational (or bureaucratic)—were interwoven at various levels of the organization structure at the Basel Mission. Discipline was, as the author shows, a crucial aspect both of the mission's structure of authority and of its perseverance over time. Yet this discipline often stood in contrast to both the potential and real expression of the charismatic impulse which is also at the heart of all evangelical activity—indeed part of its very definition from the period of late antiquity on. Both, moreover, stood in tension with the more bureaucratic formulas and legal-contractual ties (between the missionaries in the field and the committee in Basel) that defined the terms of service, the woof and weave of the organizational flow of life.

The author is at his best in these sections when describing the empirical reality, the myriad forms of mutual control and surveillance (most especially in the practice of "mutual watching" dictated by the committee

and carried out by the missionaries), the restriction of freedom in the life of the missionaries, the nature of the formal ties between them and the committee (ties of almost total subservience to a patriarchal authority), and so on. His analysis of the contradictions between the different forms of authority is restricted, however, mostly to a study of how certain individuals reacted in innovative ways to challenges in the field only to meet the approbation of the ruling committee. The author's study of the organizational tension between deviance and creativity, as exemplified in the lives of three individuals, is itself a fine microstudy of the type of contradiction that plagues tightly structured organizational frameworks. Equally exemplary is his study of how the terms of solidarity defined by the practice of mutual surveillance tended at times to erode the basis of solidarity itself. In themselves, however, these studies do not give us more generalized (and generalizable) insights into the structural contradictions that inhere in the interweaving of different forms of authority in one organizational framework. This is perhaps a necessary trade-off, given the rich ethnographic detail of the analysis. In sum, the author successfully integrates an organizational analysis with a deep sensitivity to both the cognitive and structural dimensions of how a particular ideal of social order was institutionalized in a given time and place.

*The Problem of Order: What Unites and Divides Society.* By Dennis H. Wrong. New York: Free Press, 1994. Pp. x + 354. \$24.95.

Satoshi Kanazawa  
Cornell University

The problem of order is one of the most important in all of the social sciences. Given that human beings are at least capable of behaving in selfish and egoistic ways, how is society possible? How can they cooperate with each other to produce social order? In many social scientists' minds, this question is inseparably associated with Hobbes, so much so that "the problem of order" and "the Hobbesian problem" are often used interchangeably. A major contribution of Dennis Wrong's *The Problem of Order: What Unites and Divides Society*, is to remove this profound question from the exclusive domain of Hobbes. Wrong discusses how Locke, Rousseau, Freud, Marx, and Parsons all wrote on this problem and suggested their solutions.

In order both to conceptually refine the problem of order and to acknowledge the contribution of others, Wrong distinguishes two versions of the problem. One, which he calls the "hobbesian/freudian problem," or the problem of socialization, relates to how individuals can overcome their asocial nature in order to cooperate with each other. The other, which Wrong calls the "marxian problem," addresses intergroup conflict. (Wrong uses the lowercase to emphasize that, although they deserve credit for properly defining the problem, their solutions were unsatisfac-



tory.) While the marxian problem of intergroup conflict assumes at least a partial solution of the hobbesian/freudian problem (there can be no intergroup conflict unless members of each group can cooperate with each other *within* the group), solutions to both are necessary for social order. Wrong correctly points out that many past writers have completely overlooked this important distinction. It is not clear, however, why Wrong equates the hobbesian problem with the problem of socialization. Hobbes's solution to the problem of order, *Leviathan*, unlike Freud's solution of proper development of superego, does not require socialization of individuals into social and moral beings. In fact, it is precisely because such socialization is difficult that *Leviathan* is necessary to guarantee order.

This book is a product of a lifetime of deep thinking and incredibly wide reading, and anyone interested in the problem of order should read it. Wrong often meanders and digresses from the main topic, but he asks some of the intellectually most stimulating questions when he does, such as, Why do individuals seek social approval? Why do we care what others think of us as long as they don't physically molest us? (p. 69).

Perhaps the most frustrating thing about this book is that, while it extensively addresses and discusses it, it nevertheless fails to answer the question posed in the book's subtitle: What unites and divides society? Wrong does not offer his own solution or theory of social order. He does offer, however, a very casual account of how order is produced: "Repeated interactions give rise to habits. They are perceived by the actors and become expectations in the sense of predictions or anticipations of behavior. Aware of what is expected by the other, each actor feels constrained to live up to the expectation, partly out of a feeling that the other will be irritated, offended, or disappointed if the expectation is not fulfilled. In short, interaction generates habits; perceived, they become reciprocal expectations; in addition to their purely predictive and anticipatory nature, sensitivity to them endows them with a constraining or even an obligatory character" (p. 48). Or as Wrong summarizes, "Social order consists of people's expectations which are, by and large, or most of the time, borne out because other people are aware of them and live up to them" (p. 57).

As Wrong himself recognizes, it would be "a great mistake to regard the above account as even approximating an adequate explanation of social order, for there is a huge gaping hole in it" (p. 59). In his mind, "What is omitted is any discussion of motivation, of the human wants, feelings, and impulses that undergird and set in motion the whole process of social interaction out of which relatively stable institutions develop" (p. 59). This question is important, but to me the largest missing link in Wrong's account of social order above is the process whereby "constraining" and "obligatory" norms thus created actually produce normative behavior. The existence of norms is different from the existence of normative behavior; not all norms are effectively enforced. And it seems naïve to believe that human beings obey norms just because other people would be "irritated, offended, and disappointed" if they did not.

Perhaps it is not fair to criticize Wrong for his failure to offer a general

theory of social order. For he, in his autobiography, specifically denounces general theories in sociology; he believes that historical knowledge is sufficient to explain the empirical world ("Imagining the Real," *Authors of Their Own Lives: Intellectual Autobiographies by Twenty American Sociologists*, edited by Bennet M. Berger [Berkeley and Los Angeles: University of California Press, 1990]). Nor is it his aim in this book to provide his own solution to the problem. "The aim of the present book is less to review comprehensively past discussions of the problem than to present a new, conceptually sharper statement of it" (p. 13). In my mind, Wrong reviews past discussions quite well, and he does indeed present a conceptually sharper statement of the problem of order. But those of us who seek *the* (or at least *a*) solution to this profound question, in the form of an empirically testable general theory of social order, must look elsewhere.

*The Frankfurt School: Its History, Theories, and Political Significance.* By Rolf Wiggershaus. Cambridge, Mass.: MIT Press, 1994. Pp. ix + 787. \$60.00.

Douglas Kellner  
*University of Texas at Austin*

Wiggershaus's monumental study of the history and intellectual trajectory of the Frankfurt school uses archival sources, interviews, and comprehensive analysis of published works to provide the most complete and up-to-date account of the development and contributions of one of the century's most influential group of scholars. Originally located within the Institute of Social Research in Frankfurt, the group—which included Max Horkheimer, T. W. Adorno, Erich Fromm, Friedrich Pollock, Herbert Marcuse, Leo Lowenthal, and their other associates—helped set the agenda for much social theory in the turbulent era from the 1930s through the 1960s. Through their influence on Jürgen Habermas and his associates, and on other social theorists and critics throughout the world, the Frankfurt school continues to be an important force in the contemporary era.

The translation of Wiggershaus's superb study, originally published in German in 1986, is therefore extremely welcome: anyone interested in contemporary social theory should read this book. Wiggershaus explicates the key positions of the Frankfurt school's critical theory of society and brilliantly traces its evolution and development. He lays bare for the first time the complex relations between various members of the group and raises serious questions concerning their positions at various times; thus the text is, in the best sense, a critical historiography.

Wiggershaus's text includes detailed discussions of the Frankfurt school's metatheory and their general theoretical positions, as well as their substantive contributions to contemporary social theory and criticism. Those unfamiliar with the group's complex trajectory will be surprised at the extent to which they contributed to methodological issues

and pioneered many empirical research strategies. The Frankfurt school was perhaps the first group to undertake, in conjunction with Paul Lazarsfeld, systematic critical inquiries into the roles of mass communications and culture in contemporary societies, and they inaugurated research into the authoritarian personality that drew heavily on Freudian theory. Members of the group carried out a series of pioneering studies of anti-Semitism and developed novel methods of group interviews in their studies of workers' attitudes during the late Weimar period and post-World War II German society. They also carried out "group experiments" in interviews that explored Germans' attitudes toward democracy in postwar Germany.

While in exile at Columbia University from 1934 through the late 1940s, these scholars from the Institute of Social Research produced penetrating theoretical studies of contemporary societies, including critical analyses of German fascism. Upon their return to Frankfurt in the 1950s, they introduced American methods of empirical research in Germany, although they criticized empiricist approaches that eschewed theoretical and critical positions in the so-called positivism debate. Rejecting both empty speculation without empirical research and blind empiricism without theory, the institute theorists attempted to blend theoretical construction with empirical inquiry. But critical theory also maintained normative perspectives, and Wiggershaus indicates how their standpoint of critique and critical strategies changed over the years.

Max Horkheimer emerges as the Frankfurt school's dominant figure in Wiggershaus's account, though the portrait is not always flattering. Horkheimer was director of the Institute for Social Research during the early 1930s, and throughout his exile in the United States he tightly controlled the publications and discussions of institute members. When Horkheimer, Adorno, and Pollock returned to Germany in the late 1940s, Horkheimer played a strong administrative role in Frankfurt University, but a less active role in the development of critical theory, a position that devolved on Adorno and later Habermas (while Marcuse continued to develop critical theory in the United States until his death in 1979).

Wiggershaus shows how Horkheimer wavered between wanting to guide an interdisciplinary research institute that was dedicated to developing a transdisciplinary social theory of the present age grounded in empirical studies of political economy, and wanting to engage in more philosophically oriented theoretical work with Adorno. Wiggershaus documents the shifts back and forth from interdisciplinary work with an empirical focus to more philosophical theorizing on the part of Adorno and Horkheimer in particular. The overall impression that emerges is that the Frankfurt school never really realized their project of developing an interdisciplinary theory of the present age and that Horkheimer and Adorno were more interested in engaging in purely theoretical projects, such as the studies that constituted *Dialectic of Enlightenment*, rather than carrying out interdisciplinary research.

By stressing the differences among its members and the variety of

positions that they took on fundamental issues, Wiggershaus also raises the question of whether the group really constituted a "school" at all. But Wiggershaus's study makes clear that the institute members produced a variety of works of lasting interest and importance. His lively portraits of the institute's key figures makes the personalities come alive, and he incisively presents the key ideas. Yet the Frankfurt school resembles not so much a "school," with a set method, doctrine, and defining positions, as it does a group of individuals, often with their own interests and agendas, that sometimes coalesced and sometimes diverged.

Wiggershaus pays close attention to the differences and debates within the group, and he carefully charts the complex developments of each individual thinker, as well as the changing relationships between the various members of the institute. *The Frankfurt School*, in conjunction with planned and forthcoming editions of the complete works of Horkheimer, Adorno, Fromm, Marcuse, and other key members of the group, thus makes possible a deeper understanding of the richness and variety of critical theory than was previously possible. And, finally, Wiggershaus's critical probing of their various positions shows that the institute never fully developed a critical theory of contemporary society, a task that remains a challenge for the present age.

*The Birth of Fascist Ideology: From Cultural Rebellion to Political Revolution.* By Zeev Sternhell, with Mario Sznajder and Maia Asheri, translated by David Maisel. Princeton, N.J.: Princeton University Press, 1994. Pp. x + 338. \$29.95.

Stephen Turner  
*University of South Florida*

When this book first appeared in French it produced a violently negative reaction that has continued unabated. The reaction came from those whose cultural attitudes have the same roots that Sternhell claims fascism had. Nevertheless, the case appears to have withstood the challenge. Sternhell's argument, amply documented by an extended examination of French and Italian writings, is that fascism's intellectual origins evolved from theoretical reflections on the practical political experience of the extreme revolutionary left. Sternhell begins with Sorel and the situation of French socialism at the turn of the century, when the bourgeois order was the picture of rude health. The hope of a revolution of the sort predicted by Marx, driven by the inexorable logic of the immiseration thesis, was no longer credible. The workers proved to be less than heroic. Instead, they were seduced by the material plentitude of the bourgeois order, and their political leaders were tamed by inclusion in the parliamentary system. Yet the rage against this order—against parliamentarism, individualism, materialism, positivism, and so forth—that was deeply rooted among the activists and intellectuals of the left was still potent.

Sorel was the first to grasp the situation. Implacably opposed to the existing order and parliamentary democracy, a theorist of the general strike and the heroic proletariat, he grasped that Marxism—obviously false as a predictive theory—was potent as a revolutionary myth, and that as a myth it ennobled those who acted heroically against the existing order. Parliamentarism, in contrast, tamed socialism, and participation in parliamentary liberal democracy turned revolutionary factions into organizations governed by the realities of organizational life. With this began a wrenching ideological reconstruction. The first stage was to reject those variants of Marxism that believed in reformism or in the orthodox mechanical story of the triumph of the proletariat. What was needed was unity of will, and the means for obtaining this was to strengthen antagonisms and to struggle against “human rights,” social Catholicism, and universal values, all of which diminished antagonism. The syndicalists were, for these intellectuals, the proper development of Marxism, for they had avoided these errors. The second stage was to add the element of the cult of vitalism, activism, and irrationalism, tied throughout to the idea of the general strike, which now took the place of the needed revolutionary myth. The idea that collective electricity was a condition of heroism and heroic values now replaced the rationalism of Marxism. With this, however, the transition was incomplete. When these theories were put into practice, the proletariat failed to live up to them. Despite some organizational successes and large-scale participation, general strikes quickly collapsed.

Vitalism demands vitality, and it became clear to the French fellow travelers of Sorel that it was the bourgeoisie, not the proletariat, that possessed the kind of vitality needed for radical change, and among the younger French bourgeois they found movements whose bellicosity, rage, and taste for heroism was to their liking, and that were at one with them as critics of the enlightenment and its works. Berth held “that positivism, which created the ‘regime of money, an essentially leveling, materialistic, and cosmopolitan regime’ had delivered up France to ‘the essence and quintessence of bourgeois materialism, the Jewish speculator and financier’” (p. 125). Reform socialism was the adoption of the worst ideas of the decadent bourgeoisie: pacifism, materialism, Malthusianism, and especially the “modern democratic state.”

Italian syndicalism was dominated intellectually by the French, and it was there, under the influence of the experience of several general strikes, that the practical organizational lessons that later were crucial to fascism were learned. During the prewar period the idea of the general strike was debated in socialist congresses and circles. The lessons were obscure. Strikes did work in the sense that they had powerful effects and involved people, and the experience of organizing these strikes was crucial to the practical political education of leaders. Mussolini was one revolutionary syndicalist who was formed by this experience, and through a slow intellectual development, came “to lose faith in the revolutionary virtues of the proletariat together with the organizations by which it was repre-

sented and in internationalism" (p. 213). The First World War ended both elements of his earlier faith; at the same time, tactical considerations pointed directly to the need for a more potent unifying myth and a practical model of political authority and the economy.

Nationalism had, in 1914 and even earlier, in Italy's Libyan adventure, proven its potency as a unifying myth. National socialism was the synthesis that seemed to Mussolini to contain the vitality to bring about radical change. He soon rejected Marxism, recognized (as the revolutionary syndicalists had already done) the permanence of capitalism, and in 1918 could see the Bolshevik Revolution as evidence of the political and economic failure of the proletariat. Production required, he now believed, the cooperation of the classes. Fascism was for him a revolutionary "third way" (quoted on p. 220) between a bankrupt Marxism and a degenerate liberalism. Many other European intellectuals, Sternhell shows, followed the same path, from a Rosa Luxemburg style of extreme revolutionary socialism to fascism and Nazism; others, notably cultural conservatives of the romantic type, who came from a related cultural rebellion, joined their cause.

This book is clearly a major work and has been received as a major challenge to the standard view of fascism, with implications for the understanding of Nazism. As history, it cuts a new path. It is too well documented and carefully thought-out to be dismissed, too radical in its implications to be comfortably assimilated. Though Mussolini's socialist origins were well known, the French intellectual background to his development was, to the extent that it was known, not understood as a direct source of fascism. Nor are the connections the fuzzy sorts of "influences" that sometimes figure in the writings of intellectual historians. Sternhell explains the development of fascist ideology as a logical, if not reasonable, response to practical political circumstances. Indeed, it becomes clear that, in their own way, the proto-fascist left had a better grip on political reality than many of their more ideologically rigid contemporaries in the socialist camp.

Although *The Birth of Fascist Ideology* is not a work of sociology, it bears on many of the clichés that sociologists have employed to account for fascism and Nazism (which are only briefly discussed, in the final chapter of the book). The hatred of the bourgeois order and contempt for unheroic bourgeois values that is such a visible theme both of Italian fascism and Nazism is, for the first time, given proper recognition as an integral, indeed central, part of the phenomenon of fascism, emotionally and ideologically. Standard sociological interpretations of fascism, in terms of such things as the status anxieties of the lower middle class, have generally ignored these elements of the appeal of fascism, suggesting that whatever the content of fascism, its real appeal was that it seemed to promise to put the proletariat in its place. One would hope that this book challenges sociologists to do something better: to answer the question of why the "bourgeois" order and its values have produced such bitter and violent antagonists.

*A Social History of Truth: Civility and Science in Seventeenth-Century England.* By Steven Shapin. Chicago: University of Chicago Press, 1994. Pp. xxxi + 483. \$29.95.

Kevin L. Brown  
*University of Chicago*

Hearing from the Dutch ambassador that water in Holland sometimes gets hard enough to walk on, the astonished king of Siam remarked, "I have believed the strange things you have told me, because I look upon you as a sober fair man, but now I am sure you lie" (p. 229). In his book *A Social History of Truth* Steven Shapin repeats this story by Locke to show that the factual status of natural phenomena depends upon the credibility of witnesses when direct experience is not at hand. Indeed, since a succession of "witnesses" such as manufacturers, technicians, and past researchers inevitably are needed to lay the groundwork for judging any empirical claims about nature, modern science itself "takes place on a field of trust" (p. 417). Although many sociologists have come to view scientific activity as socially embedded, Shapin more explicitly asks how the "networks of justified trust" (p. 205) that structured 17th-century English society were used to put order in the scientific revolution as well. "Who gets to speak? Whose speech may be accounted truth? These are political matters," writes Shapin, "and as they are resolved so political order is constituted, both in the house of knowledge and in the state" (p. 403).

The person who "spoke" the most for the new philosophy was the great scientist Robert Boyle. Taking Boyle as his paradigm, Shapin shows that the honor culture of the English gentry combined with Puritan notions of piety, civility, and virtue to solve the problem of order in scientific society by creating a recognizable member of the "culture of veracity" (p. 78): the Christian virtuoso. Such a person could be trusted to tell the truth, because the Christian had a nobility of spirit that the mere gentleman lacked, while the landed gentry, unlike the servant, the merchant, or even the minister, were free to act without institutional constraints. In contrast to traditional philosophers or quarrelsome scholars, then, the Christian virtuoso deployed "emblems of disinterestedness" (p. 175) in his presentation of self both to fellow members of the Royal Society and to society at large. "Having no commitments to corporations of experts and institutionalized bodies of culture, he was free to commit himself to truth. A selfless self was a free actor in the world of knowledge; all others counted as constrained" (p. 182).

Thus Shapin finds that when "free actors" offered accounts of icebergs that conflicted with basic principles of hydrostatics, Boyle blamed faulty instruments or considered alternative hypotheses, but when vulgar divers reported that water did not exert pressure at great depths, they simply were not to be believed. Similarly, Leeuwenhoek's incredible observations with his microscope needed credible witnesses and the ingenious

Robert Hooke was forced to repeat experiments in front of members of the Royal Society simply because Leeuwenhoek and Boyle belonged to the wrong social class. Servants might be needed to run experiments, but their observations were not always to be trusted, and a gentleman necessarily had to take credit for their good work. When two reliable astronomers argued over the path of a comet, the Royal Society simply posited the existence of two comets in order to "acquit the requirements of both astronomical plausibility and philosophical civility" (p. 277). In short, Shapin finds that "the norms of everyday civil conversation specified that gentlemen *were to count* as transparent spokesmen for reality" (p. 191; Shapin's emphasis).

Less transparent, however, is Shapin's finding that the norm of gentlemanly conversation that discouraged "giving the lie" at all costs entailed a probabilistic rather than a mathematical basis to the new philosophy. Boyle may not have demanded a mathematical certainty for experimental results in order to preserve civility but simply because, as Shapin observes, physical and technical constraints made these results inherently uncertain. And as Shapin also observes, expectations based on technical grounds are not *moral* norms since they do not involve the promised actions or communications of people (pp. 7–8). This distinction itself may seem rather technical, but many sociologists similarly have blurred the line between the social embeddedness of science and its nonsocial content. Perhaps what is most remarkable about Shapin's "historical ethnography" (p. xvi) is that it shows that the first "moderns" did attempt to make accurate observations, did attempt to minimize their dependence on second-hand information, and did attempt to maximize the correlation between reliability and trust. By finely demarcating the role that society played in this revolutionary task, *A Social History of Truth* has challenged these sociologists to explain how early scientists discovered many facts about nature or, more importantly, to explain why, unlike the king of Siam, they even bothered to look.

*Durkheim and Women.* By Jennifer M. Lehmann. Lincoln: University of Nebraska Press, 1994. Pp. x + 173. \$30.00.

Edward A. Tiryakian  
*Duke University*

"Socialism," said Durkheim in 1896, "is not a science. . . . It is a cry of pain and sometime of anger, uttered by men who feel deeply our collective malaise" (*Le socialisme* [Paris: PUF, 1971], p. 37). Jennifer Lehmann's feminist version of socialism, "critical structuralism," whose launching pad is a radical critique of Durkheim's treatment of women, will strike some as a cry of fury by a latter-day Erinye. The rationale for pursuing Durkheim the length of the book, I infer, is that his social determinism might have been used to promote feminism, but instead this



sociological Orestes commits the heinous deed of opting for "capitalist patriarchy," as "he denies women humanity and condemns them eternally to the netherworld of primitive, animal, physical, biological structures" (p. 30). Durkheim, we are told, had a twin fear of socialism and feminism (p. 31), hence his inability to take his social determinism beyond "the totalizing, pervasive and tenacious liberal ideology of bourgeois capitalist, patriarchal hegemony" (p. 125).

The sweeping critique of Durkheim does make the telling point (not altogether original with the author) that Durkheim's general analysis of modern society is flawed because he tacitly relegates women primarily to the domestic sphere and subordinates their interest to that of men. This comes out in his relevant pieces *Suicide* and "Divorce by Mutual Consent," where he admits that marriage is beneficial to men but not to women as a preventive of suicide. Divorce and widowhood have a more negative impact on men than women (using suicide rates as indicators). Durkheim's functionalism in effect equated "society" with the public sphere, in his time essentially dominated by men as the fully socialized products of society. Hence came his apparent willingness to "sacrifice" women to marital confines for the greater good of society but really of men who might otherwise fall into the trap of sexual anomie.

Although Lehmann undeniably has a rich polemical repertoire (p. 133), this work hurls so many shibboleths that scholarship seems secondary. Lehmann's treatment of Durkheim's milieu (chap. 1) is grossly inadequate. She mentions a few persons of the French Revolution and early 19th century who spoke on behalf of women's equality (she misses an opportunity in not discussing the Saint-Simonians), but she gives no scholarly analysis of the situation of women and the feminist movement in Durkheim's own historical period in France and other Western industrial countries. Durkheim's views of women might be judged as patriarchal by today's "evolved" standards, but they were common currency in the Victorian age, among liberals as well as conservatives (and it might be borne in mind that even in groups with a radical ideology of equality such as SDS in the late 1960s, the exploitation of women followers was a significant factor in the emergence of the women's movement in the 1970s). In order to prove her point, Lehmann would need to present an extensive historical comparative examination of the relation of men to women in the public sphere, the public debates regarding divorce by mutual consent, and the issue of women's suffrage and to discuss the perspective of Durkheim relative to fin-de-siècle sociologists, socialists, and other intellectuals (such as the psychoanalytical school) on both sides of the Atlantic.

Moreover, Lehmann would have done well to discuss carefully the full range of texts in which Durkheim deals with women, sexuality, and divorce, such as those brought together by Victor Karady (*Émile Durkheim, Textes*, vol. 2 [Paris: Editions de Minuit, 1975]), including a notable public debate on divorce in 1909 sponsored by L'Union pour la Vérité in which Durkheim was challenged by a Mme. Compain and a

Mlle. Chambon. A little research on these two hardy women might have added important concrete materials to a volume entitled *Durkheim on Women*.

The book is further marred by Lehmann's loose use of terms. For instance, she tars Durkheim's "organicism" (p. 75), but her discussion of this topic is more appropriate to Spencer or the young René Worms than to Durkheim. Lehmann repeatedly makes Durkheim an exponent of capitalism and patriarchy, but I would like to see citations where Durkheim extols capitalism (a critic of unregulated capitalism, Durkheim even advocated the abolition of inherited wealth). The patriarchy he discussed (Max Weber did this more fully in *Economy and Society*, vol. 2 [Berkeley: University of California Press, 1978], chap. 12) was the historical form of the Roman family, which Durkheim in his book reviews found to have improved the situation of women from preceding stages of family organization. Patriarchy has no unique linkage to either feudalism or to capitalism.

Lehmann missed an opportunity here to deploy her genuine theoretical bent in a more dispassionate discussion of the relevance of Durkheim and feminist sociological theory. She unnecessarily repeats herself in chapters 3 and 4, and the last chapter ("The Contemporary Milieu") is a cursory overview of highly abstract current ideologies and discourses. Frank Pearce's *The Radical Durkheim* (London: Unwin Hyman, 1989) and Miriam Johnson's "Functionalism and Feminism: Is Estrangement Necessary?" (in *Theory on Gender/Feminism on Theory*, edited by Paula England [New York: Aldine, 1993]) are more effective dialogues between radicals and mainstream sociological positions, while on the other hand, "critical structuralism" seems left behind by more recent radical sociological perspectives on sexuality (see, e.g., the July 1994 issue of *Sociological Theory*, esp. Chrys Ingraham's "The Heterosexual Imagery: Feminist Sociology and Theories of Gender").

There are projects that might, in the spirit of the Eumenides, fruitfully combine a commitment to feminism and an interesting ambivalence toward Durkheim. One might be a historical study of feminists, feminism, and the emergent social sciences at the turn of the century in both Europe and America. A second would be a Durkheimian analysis of contemporary feminism and its social basis in the vein of his monumental but uncompleted study of socialism. A third might be a Durkheimian-feminist upgrade of *Suicide* and *The Division of Labor*. Such empirical studies could serve as documentation for a more theoretical work in the nature of a dialogue between feminist theory and Durkheimian sociology, which Lehmann herself intimated is possible in stating "Durkheim's scientific theory of social determinism could be appropriated by feminism as an alternative to the 'scientific' theory of individualism" (p. 125). Invectives aside, Lehmann has provided us with a challenging new reading of Durkheim, and the jolts her reading produces should be seen as salubrious for new ventures in sociological analysis.

## Book Notes

*Democratic Reform and the Position of Women in Transitional Economies*. Edited by Valentine M. Moghadam. Oxford: Clarendon Press, 1993. Pp. viii + 366. \$60.00.

*Democratic Reform and the Position of Women in Transitional Economies* examines how the transition from communism to capitalism has affected the position of women in Eastern Europe and the former Soviet Union. The 14 articles in the collection examine issues such as the decrease in political participation of women, educational attainment of women, attitudes toward gender roles, and labor market experiences of women in the context of dramatic social change. Not surprisingly, the authors of this volume agree that women have disproportionately suffered in the shift of political and economic regimes. These patterns of increasing gender inequality are evidenced by, for instance, the high unemployment rate of women in the former East Germany that has resulted from unification and privatization. Specifically, quotas that promoted the training and participation of women in political structures and the labor market have been abolished in the post-Communist era. Some of the authors attribute the decline of the status of women to the uneven promotion of women's rights in the socialist era, thus making these issues susceptible to a growing conservative movement in many countries. Contributors include the editor, Marilyn Rueschemeyer, Gillian Paull, and others. This volume will be valuable to students and researchers interested in descriptive accounts of how recent changes in Eastern Europe and the former Soviet Union have affected the status of women.

*After Ellis Island: Newcomers and Natives in the 1910 Census*. Edited by Susan Cotts Watkins. New York: Russell Sage Foundation, 1994. Pp. xviii + 451. \$49.95.

The 1910 census provides a glimpse of the extensive flow of immigrants that began in the 1880s and was quickly stemmed by World War I and the onset of restrictive immigration statutes. This particular decennial census is also unique because it includes detailed questions about national origin and, for the first time, items about mother tongue, thus allowing for detailed classifications of ethnic groups. *After Ellis Island* addresses myriad aspects of the immigrant experience—specifically, mortality, fertility, families and households, neighborhoods, schooling, and industrial

Book notes are written by members of the book review editorial board.  
© 1995 by The University of Chicago. All rights reserved.  
0002-9602/95/10005-0008\$1.50

affiliation—and how these experiences differ by regional groups. Watkins emphasizes that the tabulations not only provide descriptive information about the population but allow an examination of the differential propensities for assimilation among various national and regional groups. Furthermore, by addressing the assimilation experience of immigrants in the early part of the 20th century, the articles in this book offer new insights into the process of adaptation of today's immigrant population. Among the contributors are Samuel Preston, S. Philip Morgan, Michael J. White, Ewa Morawska, Jerry Jacobs, Ann Miller, and Michael Strong. Although these essays will not resolve current debates about the nature of the assimilation process, this volume provides an invaluable reference for researchers interested in the early experiences of immigrants to the United States.

*Chinese Families in the Post-Mao Era.* Edited by Deborah Davis and Stevan Harrell. Berkeley and Los Angeles: University of California Press, 1993. Pp. xii + 370. \$55.00 (cloth); \$17.00 (paper).

This volume uses field research and survey data from urban and rural areas to focus on the impact of post-Mao reforms on family life in China. Specifically, the authors are concerned with how Chinese families have responded to the easing of strict policy controls over marital practices, fertility, and the care of the elderly. Besides Davis and Harrell, who provide discrete chapters and an excellent introduction and summary of how modern Chinese families respond to economic and political stimuli, the book's contributors include Martin King Whyte and Susan Greenhalgh, who discuss wedding behavior and the one-child policy. Their findings suggest that, in contrast to assertions of earlier studies about the secondary role of government policies in determining family life in the late-Mao era, coercive public policy was a major force in reforming marriage practices. In other words, changes in family life may have been less a consequence than a cause of economic development. In fact, recent years have witnessed a resurgence of certain rituals and behavior that not only emphasizes the importance of policy in restricting such activities in Mao's lifetime but also suggests a response to the political economies of the new era. Examinations of these new forms of family activities hint at new definitions of kinship ties. Finally, the authors find heterogeneity in family life in post-Mao China but conclude that Chinese families generally adapt their home life strategically to local economic conditions. Overall, this is an important volume that will interest not only scholars of China but also those who study the links between social change and family structure.

*The Authority of the Consumer.* Edited by Russell Keat, Nigel Whiteley, and Nicholas Abercrombie. New York: Routledge, 1994. Pp. x + 281. \$69.95 (cloth); \$22.50 (paper).

This collection of essays addresses an interesting theoretical issue: Does the advent of consumer society bring about a shift in authority from producer/provider to consumer/recipient? Is authority a useful analytic concept when applied to the producer/consumer relationship? The first of three sections is theoretical, including a lengthy introduction by the editors. The second and third sections cover, respectively, cultural institutions and the provision of public services, both from a consumption perspective. This is a somewhat unusual volume that includes contributions from professionals working in the arts, advertising, and public service as well as academics, all of whom are associated in some way with the Centre for the Study of Cultural Values at Lancaster University. The book is provocative and deals with important issues, with the chapters by Abercrombie, Alan Warde, and Baz Kershaw among the best. This volume will be of interest to anyone studying patterns of consumption and the structure of producer/consumer relations in modern societies.

*Cultures in Flux: Lower-Class Values, Practices, and Resistance in Late Imperial Russia.* Edited by Stephen P. Frank and Mark D. Steinberg. Princeton, N.J.: Princeton University Press, 1994. Pp. ix + 214. \$49.50 (cloth); \$14.95 (paper).

This volume contains several descriptions of popular culture in turn-of-the-century Russia. From death rituals to folk songs to amusement parks to the penny press, the chapters are well written and describe a wide range of popular cultural forms and practices. In some cases there is not much explicit theoretical elaboration. However, the volume is important because of how the descriptions themselves are written. The authors problematize resistance and question such unitary categories as "peasant" and "worker." In doing so, they introduce a subtlety not often found in accounts of popular culture, which tend to stress either the unproblematic reproduction of dominant ideologies or find radical resistance in each utterance of a marginalized group. This volume is a good introduction to historical work on popular culture, power, identity, and resistance and will be of interest to sociologists interested in any of these areas.

*Workplace Industrial Relations and the Global Challenge.* Edited by Jacques Belanger, P. K. Edwards, and Larry Haiven. Ithaca, N.Y.: ILR Press, 1994. Pp. ix + 325. \$58.00 (cloth); \$26.95 (paper).

This edited volume is a welcome addition to increasingly important, but rarely done, studies of workplaces: ethnographic and comparative studies of labor regulation in a global context. Previous studies tend to be either

rich descriptions of the shop floor that ignore differences between societies or international comparisons that do not acknowledge variations within a society. This volume seeks to combine the two traditions.

The volume consists of nine chapters dealing with not only countries that have been studied extensively (such as the United States, Canada, Mexico, Australia, and Germany), but also East Asian countries (such as Singapore). Throughout the chapters, the key issue is "the way in which the workplace is regulated in different countries, together with the extent of variation within each country and the reasons for and results of different methods of regulation" (p. 9).

Methodologically, this volume has three forms of inquiry: first, a detailed ethnographic analysis of the workplace with observation and interviews; second, a historical analysis of the negotiation of control on the shop floor; finally, more macrolevel comparison of workplaces in different countries.

Substantively, this book deals with four areas: comparisons among the Anglo-Saxon countries of the North Atlantic; national systems versus workplace peculiarity; the influence of U.S. capital; and globalization. The first area concerns the extent and conditions of the different development of workplace relations in Britain and North America as well as the divergence in labor relations between the United States and Canada. The second theme attempts to demonstrate that national systems are neither homogeneous entities nor sets of disparate cases. The contributors show that there are constant interactions between national and workplace levels, producing variations within and between societies. The third theme focuses on how American multinational firms affect and are affected by the nature of workplaces peculiar to different countries. The last theme concerns how workplaces in each country deal with the globalization of labor markets and how labor regulations are changing accordingly.

This book is a valuable contribution to the study of industrial relations methodologically as well as substantively. It also illustrates the benefits to be reaped from combining ethnographic and comparative methods.

# SOCIOLOGY CALIFORNIA

Charles M. Payne  
**I've Got the  
Light of  
Freedom**

The Organizing Tradition and the Mississippi Freedom Struggle

"There are invaluable lessons here on how social change can actually occur.

... We learn that within each of us resides the levers of social change only waiting to be activated."

—Aldon Morris,  
author of *The Origins of the  
Civil Rights Movement*

*A Centennial Book*

\$28.00 cloth, illustrated

Tomás Almaguer  
**Racial Fault Lines**

The Historical Origins of  
White Supremacy in  
California

"The last and best word on the historical origins of the racial hierarchy that contemporary multiculturalists are struggling to overcome."

—George Fredrickson,  
Stanford University

\$40.00 cloth, \$15.00 paper

At bookstores or order toll-free  
1-800-822-6657



Richard Madsen  
**China and the  
American Dream**

*A Moral Inquiry*

Madsen's frank and innovative examination of the moral history of U.S.-China relations targets the forces that have shaped this surprisingly strong tie between two strikingly different nations.

*A Philip E. Lilienthal Book,*  
\$27.50 cloth

Craig Calhoun  
**Neither Gods nor  
Emperors**

*Students and the Struggle  
for Democracy in China*

In this eyewitness examination of China's 1989 student protest movement, Calhoun offers a dramatic look at the momentous uprising which captured the world's attention.

\$37.50 cloth

From *I've Got the Light of Freedom*

## UNIVERSITY OF CALIFORNIA PRESS

# SOCIOLOGY CALIFORNIA

Ellen Herman

## **The Romance of American Psychology**

Political Culture in the Age of Experts

"A wonderfully written book . . . [about] a little-recognized but enormously significant process that has shaped contemporary American political culture." —Cynthia Enloe,

author of

*The Morning After*

\$35.00 cloth, illustrated

Mary-Jo DeVecchio  
Good, Paul E. Brodwin,  
Byron J. Good, & Arthur  
Kleinman, Editors

## **Pain as Human Experience**

An Anthropological Perspective

"A sophisticated and engaging book that takes us into subtle phenomenal worlds that ordinarily escape the attention of social scientists."

—*American Journal of Sociology*  
*Comparative Studies of Health Systems and Medical Care*, \$14.00 paper

At bookstores or order toll-free  
1-800-822-6657.

Christine L. Williams

## **Still a Man's World**

Men Who Do  
"Women's Work"

In this eye-opening study, Williams provides a wholly new look at men who work in predominantly female jobs.

*Men and Masculinity*, \$40.00 cloth,  
\$15.00 paper



Jane Schneider and  
Rayna Rapp, Editors

## **Articulating Hidden Histories**

Exploring the Influence of  
Eric R. Wolf

Here many scholars influenced by Eric R. Wolf acknowledge the contribution of this great scholar while extending his work by presenting their own original field and archival research.

\$45.00 cloth, \$18.00 paper

From *The Romance of American Psychology*

# UNIVERSITY OF CALIFORNIA PRESS



# Population Research and Policy Review

In cooperation with the Southern Demographic Association (SDA)

Editor **David F. Sly**

Associate Editors **Isaac Eberstein,**  
**Carl Schmertman, William Serow, E. Walter Terrie**

Editor **Charles Nam** (Book Review Editor)

Managing Editor **Cherise Davis**

*Population Research and Policy Review* has a two-fold goal: it provides a convenient source for government officials and scholars in which they can learn about the policy implications of recent research relevant to the causes and consequences of changing population size and composition; and it provides a broad, interdisciplinary coverage of population research.

*Population Research and Policy Review* seeks to publish quality material of interest to professionals working in the fields of population, and those fields which intersect and overlap with population studies. The publication includes demographic, economic, social, political and health research papers and related contributions which are based on either the direct scientific evaluation of particular policies or programs, or general contributions intended to advance knowledge that informs policy and program development.

*Population Research and Policy Review* is abstracted and/or indexed in: *Current Contents/Social & Behavioral Sciences*, *Current Index to Journal in Education*, *Economic Literature Index*, *Family Abstracts*, *Family Resources Database*, *Family Studies Abstracts*, *Geo Abstracts*, *Index to Periodical Articles Related to Law*, *International Political Science Abstracts*, *Journal of Regional Abstracts*, *Pais Bulletin*, *Population Index*, *Review of Population Reviews*, *Sage Public Administration Abstracts*, *Sage Urban Studies Abstracts*, *Social Sciences Citation Index*, *Social Work Research and Abstracts*, *Journal of Economic Literature* and *The Survey of Regional Literature*, *The International Bibliography of the Social Sciences* and *Cambridge Scientific Abstracts*, *Sociological Abstracts*, *Academic Abstracts*, *PAIS International*, *Research Alert*

## Subscription Information

ISSN 0167-5923

1994, Volume 13 (4 issues)

Subscription rate: Dfl. 398.00/US\$208.00 *Incl. postage and handling*

PO Box 322, 3300 AH Dordrecht, The Netherlands  
PO Box 358, Accord Station, Hingham, MA 02018-0358, USA

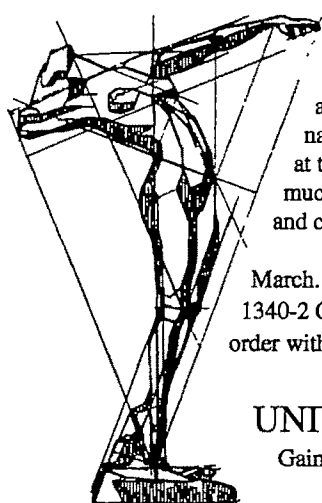
Journal  
Highlight

KLUWER  
ACADEMIC  
PUBLISHERS



# Social Science and the Challenge of Relativism, Volume III. Cultures of Nature:

## An Essay on the Production of Nature



Lawrence Hazelrigg

"A quality piece of work; the central problematic is clearly articulated and important; the theoretical analyses are sophisticated and subtle; and the narrative is well crafted. . . . The focus of this work is at the heart of core issues now being discussed by much larger circles of interdisciplinary social theorists and cultural studies scholars."

—Robert Antonio, University of Kansas

March. 368 pp. Notes, references, index. ISBN 0-8130-1340-2 Cloth, \$59.95. Through full-service bookstores or order with VISA or M/C toll free: 1-800-226-3822.

*Celebrating fifty years!*

**UNIVERSITY PRESS OF FLORIDA**

Gainesville, Tallahassee, Tampa, Boca Raton, Pensacola,  
Orlando, Miami, Jacksonville

## BEYOND THE LAW

### Crime in Complex Organizations

Edited by Michael Tonry and Albert J. Reiss, Jr.

Volume 18 in the *Crime and Justice* series focuses on the latest research and commentary on criminal misconduct occurring in U.S. corporations and government agencies, with special emphasis on savings and loan fraud, insider trading, nursing homes, environmental pollution, and other forms of "white collar crime."

Leading experts from a variety of disciplines investigate these antisocial and unlawful behaviors, and further explore the roles of self-regulation, administrative regulation, and civil and criminal sanctions in shaping organizational behavior.

1993	320 pp.	Cloth	\$35.00	ISBN: 0-226-80821-1
		Paper	\$19.95	ISBN: 0-226-80823-8

Order through your bookstore or directly from  
**The University of Chicago Press**  
11030 S. Langley Ave., Chicago, IL 60628  
Or call: (312) 568-1550

6/94



THIS  
PUBLICATION  
AVAILABLE  
FROM UMI

This publication is available from UMI in one or more of the following formats:

- In Microform—from our collection of over 18,000 periodicals and 7,000 newspapers
- In Paper—by the article or full issues through UMI Article Clearinghouse
- Electronically, on CD-ROM, online, and/or magnetic tape—a broad range of ProQuest databases available, including abstract-and-index, ASCII full-text, and innovative full-image format

Call toll-free 800-521-0600, ext. 2888, for more information, or fill out the coupon below.

Name \_\_\_\_\_

Title \_\_\_\_\_

Company/Institution \_\_\_\_\_

Address \_\_\_\_\_

City/State/Zip \_\_\_\_\_

Phone (     ) \_\_\_\_\_

I'm interested in the following title(s) \_\_\_\_\_

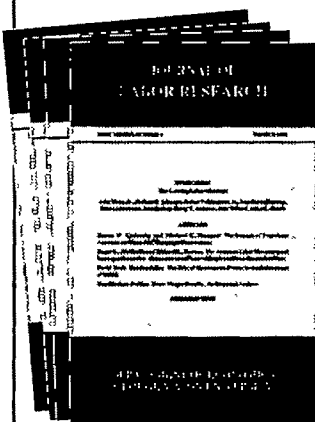
\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

UMI  
A Bell & Howell Company  
Box 78  
300 North Zeeb Road  
Ann Arbor, MI 48106  
800-521-0600 toll-free  
313-761-1203 fax

UMI



# JOURNAL OF LABOR RESEARCH

## Selected Articles Forthcoming in Volume XVI, 1995

<b>Unit Costs, Legal Shocks, and Unionization in Construction</b>	Steven G. Allen
<b>Unemployment, Union Density, and Wages</b>	Terry Ashley and Ethel B. Jones
<b>The Internal Union Political Imperative for UAW Pattern Bargaining</b>	John W. Budd
<b>Member Support for Union Mergers: An Analysis of an Affiliation Referendum</b>	John A. McClendon, Jill Kriesky, and Adrienne Eaton
<b>Marginal Revenue and Labor Strife in Major League Baseball</b>	Dale R. Oorlog
<b>Self-Interests and Group-Interests in Employee Involvement Programs: A Case Study</b>	Denis Collins
<b>Product Selection and Costs in a Partially Unionized Industry</b>	Wesley W. Wilson, Merwin C. Mitchell, and Joe A. Stone
<b>Protective Service Unions' Political Activities and Departmental Expenditures</b>	Rafael Gely and Timothy D. Chandler
<b>Public Transit Wage Rates: Pre-Reagan and Reagan-Bush Eras</b>	Ann Schwarz-Miller and Wayne K. Talley
<b>Right-to-Work Laws and Union Density: New Evidence from Micro Data</b>	Joe C. Davis and John H. Huston
<b>Union Influence on Plant Size</b>	Martin Milkman and Merwin Mitchell
<b>The Ideologies of U.S. and Canadian IR Scholars: A Comparative Analysis and Construct Validation</b>	John Godard
<b>State and Local Government Wage Differentials: An Intrastate Analysis</b>	Dale Belman and John S. Heywood
<b>Dual Commitment: Unique Construct or Epiphenomenon?</b>	Brian Bemmels
<b>Empirical Evidence on the Union Free-Rider Problem: Do Right-to-Work Laws Matter?</b>	Russell S. Sobel
<b>The Changing Character of Strikes Involving Permanent Striker Replacements, 1935-1990</b>	Michael H. LeRoy

The *Journal of Labor Research* is published quarterly by the Department of Economics of George Mason University. Subscriptions are on a calendar year basis and subscribers will receive all issues for the current year. Annual subscription rates are: individuals, \$35.00; academic libraries, institutions, government agencies, and business firms, \$85.00. Foreign subscribers add \$5.00 to cover additional postage. All remittances should be made payable to *Journal of Labor Research*, Department of Economics, George Mason University, Fairfax, Virginia 22030.

Editorial correspondence and manuscripts should be sent to James T. Bennett, Editor, *Journal of Labor Research*, Department of Economics, George Mason University, Fairfax, VA 22030. Manuscripts should be submitted in triplicate with a 100-word abstract. A submission fee of \$20.00 for nonsubscribers must accompany all manuscripts. Papers accepted for publication must conform to JLR style requirements; style sheet available on request.

# WESTVIEWPRESS

## Ain't No Makin' It

Delayed Aspirations in a Low-Income  
Suburban School

Edited by Edith Kurzweil

Jay MacLeod

The first in the series of Urban and Suburban and the story is introduced into the social and cultural world of a low-income urban school.

July 1995 • 338 pages • \$55.00 hc • \$15.95 pb

## Rethinking the Subject

An Anthropological and Literary  
Approach to Social Thought

edited by James D. Faulstich

A splendid collection of essays by leading contemporary European social theorists. Faulstich, an American, provides a book opens up a new world of thought and culture but even more importantly, it is the best."

Robert S. Bellah,

Professor of Sociology, University of California, Berkeley

227 pages • \$55.00 hc • \$16.95 pb

## Social Stratification

Class, Race, and Gender  
in Sociological Perspective

edited by David B. Grusky

A new important collection of articles on a significant and controversial topic. A collection of essays by leading sociologists and anthropologists on the most recent work in the field.

Alvin G. Sorenson,  
University of California, Berkeley

270 pages • \$69.50 hc • \$29.95 pb

See all the quality series

## The New Role of Women

Family Formation in Modern Society

Hans-Peter Blossfeld

April 1995 • 272 pages • \$55.00

See all the quality series

## Freudians and Feminists

Edith Kurzweil

Edith Kurzweil's thoughtful and occasionally technical exploration of the complex relationships between Freudians and Feminists, interestingly illuminates the ways in which both movements reflect and shape our complex times."

Elisabeth Fox-Genovese,

Emory University

202 pages • \$55.00 hc • \$17.95 pb

See all the quality series

## Social Theory

The Multicultural and Classic Reading

edited by Charles Lemert

A rich, highly textured, historically sweeping, and strikingly in these cold, that it aims to construct perhaps for the first time, a critical dialogue of contemporary social thought.

Jeffrey Alexander

672 pages • \$68.50 hc • \$23.95 pb

## Rape and Society

Readings on the Problem  
of Sexual Assault

edited by Patricia Searles

and Ronald Berger

Intellectual and sophisticated articles, narratives, and poems which explore the complex of issues in this area and present the best research in the field for teaching.

Eleanor M. Miller,

University of Wisconsin, Milwaukee

298 pages • \$55.00 hc • \$19.95 pb

## Westview Press

5300 Central Avenue

Boulder, CO • 80301-2037

Tel: (303) 441-3541 • Fax: (303) 440-3456

1-800-456-1995, orders only please

0002-9602(199503)100:5;1-6



0002-9602(199503)100:5;1-6

# AJS

American

Journal

Sociology

Volume 100 Number 6

May 1995

*Constitutional Revolutions in the Ottoman Empire, Iran, and  
Russia, 1905–1908——Sohrabi*

*Shifting Currents in Critical Sociology of Education——Davies*

*Strikebreaking, Solidarity: The Great Steel Strike of  
1919——Brown and Boswell*

## *SYMPOSIUM: PREDICTION IN THE SOCIAL SCIENCES*

*Introduction——Hechter*

*Inevitability of Future Revolutionary Surprises——Kuran*

*The Case of the Soviet Collapse——Collins*

*To Explain Political Processes——Tilly*

*And Comments by Coleman, Kiser, and Portes*

The University of Chicago Press

MARTA TIENDA, Editor

JIM DAVIS, FRED KNISS, SUSAN E. MAYER, JEFFREY MORENOFF, EMILIO PARRADO, and ROBERT J. SAMPSON, Associate Editors

MARY C. BRINTON, Associate Editor, Book Review

GRACE KAO, Assistant Book Review Editor

CHAN-UNG PARK, STEPHEN B. PLANK, MICHAEL J. REAY, DAVID A. SMILDE, and ALFORD A. YOUNG, JR., Book Review Editorial Board

SUSAN ALLAN, Managing Editor

LAUREN FORD-TAGGART and CLAIRE JAMES, Editorial Assistants

**Consulting Editors** PAUL D. ALLISON • EDWIN AMENTA • PETER BEARMAN • ROBERT J. BURSİK, JR. • JOHN CASTERLINE • NANCY DAVIS • NANCY DENTON • CLAUDE FISCHER • ROBERTO FRANZOSI • ADAM GAMORAN • JOHN R. HALL • MICHAEL T. HANNAN • MICHAEL HECHTER • DOUGLAS HECKATHORN • BARBARA HEYNS • CHARLES HIRSCHMAN • K. JILL KIECOLT • TIMUR KURAN • GARY LAFREE • DAVID LAITIN • LINDA D. MOLM • DANIEL MONTI • JEYLAN T. MORTIMER • DANIEL OLSON • DUDLEY POSTON • CECILIA RIDGEWAY • CATHERINE ROSS • RUBEN G. RUMBAUT • JOACHIM J. SAVELSBERG • DARREN E. SHERKAT • SHELLY SMITH • C. MATTHEW SNIPP • CHARLOTTE STEEH • GILLIAN STEVENS • ROBIN STRYKER • EDWARD TELLES • YU XIE

**International Consulting Editors** JOHN FOX (Canada) • MATTHIJS KALMIJN (The Netherlands) • NOAH LEWIN-EPSTEIN (Israel) • STEVEN L. RYTINA (Canada) • YUVAL YONAY (Israel)

**Editorial Board** ANDREW D. ABBOTT • GARY S. BECKER • CHARLES E. BIDWELL • DONALD J. BOGUE • RONALD BURT • TERRY N. CLARK • JAMES S. COLEMAN • JOHN L. COMAROFF • ROGER V. GOULD • WENDY GRISWOLD • EVELYN KITAGAWA • EDWARD O. LAUMANN • DONALD N. LEVINE • ALBERTO PALLONI • WILLIAM L. PARISH • MOISHE POSTONE • MARTIN RIESEBRODT • GEORGE P. STEINMETZ • ROSS M. STOLZENBERG • FRED L. STRODTBECK • GERALD D. SUTTLES • LINDA J. WAITE • WILLIAM JULIUS WILSON • KAZUO YAMAGUCHI

THE AMERICAN JOURNAL OF SOCIOLOGY (ISSN 0002-9602) is published bimonthly in July, September, November, January, March, and May by The University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, Illinois 60637. Subscription rates, U.S.A.: institutions, 1 year \$96.00, individuals, 1 year \$41.00. Student subscription rate, USA: 1 year \$28.00 (photocopy of valid student ID must accompany subscription). ASA and BSA members, 1 year \$36.00. All other countries add \$9.00 for each year's subscription to cover postage. Canadian subscribers add 7% GST. Subscription agent for Japan: Kinokuniya Company, Ltd. Single copy rates: institutions \$16.00, individuals \$6.85. Volumes 1-90 available from Schmidt Periodicals GmbH, Dettendorf, D-83075 Bad Feilnbach, Germany. All volumes are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106; and Kraus Reprint and Periodicals, Route 100, Millwood, New York 10546. Subscriptions are payable in advance. Subscriptions will be entered to start with the first issue to be published after receipt of order. **Postmaster:** Send address changes to *American Journal of Sociology*, The University of Chicago Press, Journals Division, P.O. Box 37005, Chicago, Illinois 60637.

ADVERTISING space in THE AMERICAN JOURNAL OF SOCIOLOGY is available, as is rental of its subscriber list. For information and rates, please contact the advertising sales staff, The University of Chicago Press, Journals Division, 5720 South Woodlawn Avenue, Chicago, Illinois 60637. Advertising and list rental is limited to material of scholarly interest to our subscribers.

**Copying beyond Fair Use** The code on the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made beyond those permitted by Section 107 or 108 of the U.S. Copyright Law provided that copies are made only for personal or internal use or for the personal or internal use of specific clients and provided that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc. Operations Center, 27 Congress St., Salem, Massachusetts 01970. *To request permission for other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale, kindly write to Permissions Department, The University of Chicago Press, 5801 S. Ellis Ave., Chicago, Illinois 60637. If no code appears on the first page of an article, permission to reprint may be obtained only from the author.*

Second-class postage paid at Chicago, Illinois, and at additional mailing offices.

© 1995 by The University of Chicago. All rights reserved.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984 (∞)

AJS

*American Journal of Sociology*

Volume 100 No. 6 May 1995

**CONTENTS**

- 1383 Historicizing Revolutions: Constitutional Revolutions in the Ottoman Empire, Iran, and Russia, 1905–1908  
NADER SOHRABI
- 1448 Leaps of Faith: Shifting Currents in Critical Sociology of Education  
SCOTT DAVIES
- 1479 Strikebreaking or Solidarity in the Great Steel Strike of 1919: A Split Labor Market, Game-Theoretic, and QCA Analysis  
CLIFF BROWN AND TERRY BOSWELL

**Symposium on Prediction in the Social Sciences**

- 1520 Introduction: Reflections on Historical Prophecy in the Social Sciences  
MICHAEL HECHTER
- 1528 The Inevitability of Future Revolutionary Surprises  
TIMUR KURAN
- 1552 Prediction in Macrosociology: The Case of the Soviet Collapse  
RANDALL COLLINS
- 1594 To Explain Political Processes  
CHARLES TILLY
- 1611 What Can Sociological Theories Predict? Comment on Collins, Kuran, and Tilly  
EDGAR KISER
- 1616 Comment on Kuran and Collins  
JAMES S. COLEMAN



- 1620 On Grand Surprises and Modest Certainties: Comment on Kuran, Collins, and Tilly  
ALEJANDRO PORTES

### Book Reviews

- 1627 *Critical and Effective Histories: Foucault's Methods and Historical Sociology* by Mitchell Dean  
BARRY SMART
- 1629 *Durkheim's Philosophy of Science and the Sociology of Knowledge: Creating an Intellectual Niche* by Warren Schmaus  
MARK A. SCHNEIDER
- 1630 *Language and Reason: A Study of Habermas's Pragmatics* by Maeve Cooke  
JAMES J. CHRISS
- 1632 *The Sociology of Deviance: An Obituary* by Colin Sumner  
SUDHIR ALLADI VENKATESH
- 1634 *Illegitimacy, Sex and Society: Northeast Scotland, 1750-1900* by Andrew Blaikie  
PETER BECKER
- 1636 *Gay New York: Gender, Urban Culture, and the Making of the Gay Male World, 1890-1940* by George Chauncey  
VERN L. BULLOUGH
- 1637 *Orpheus and Power: The "Movimento Negro" of Rio de Janeiro and São Paulo, Brazil, 1945-1988* by Michael George Hanchard  
RICHARD GRAHAM
- 1640 *Urban Unrest in the Middle East: A Comparative Study of Informal Networks in Egypt, Iran, and Lebanon* by Guilain Denoeux  
NADER SOHRABI
- 1642 *The Policy Challenge of Ethnic Diversity: Immigrant Politics in France and Switzerland* by Patrick Ireland  
CARL-ULRIK SCHIERUP
- 1644 *The Political Ecology of the Modern Peasant: Calculation and Community* by Leslie E. Anderson  
MARC EDELMAN
- 1646 *Gender and Power in Rural North China* by Ellen R. Judd  
KENNETH POMERANZ

- 1648 *Childerley: Nature and Morality in a Country Village* by Michael Mayerfeld Bell  
WENDY GRISWOLD
- 1651 *Demography of Aging*, edited by Linda G. Martin and Samuel H. Preston  
IŞIK A. AYTAÇ
- 1653 *Old Age: Constructions and Deconstructions* by Haim Hazan  
ANNE FONER
- 1655 *Succeeding Generations: On the Effects of Investments in Children* by Robert H. Haveman and Barbara L. Wolfe  
SUSAN E. MAYER
- 1657 *Beyond Adolescence: Problem Behavior and Young Adult Development* by Richard Jessor, John Edward Donovan, and Frances Marie Costa  
ROBERT B. CAIRNS
- 1660 *Parents' Jobs and Children's Lives* by Toby L. Parcel and Elizabeth G. Menaghan  
JANET CURRIE
- 1662 *Flesh Peddlers and Warm Bodies: The Temporary Help Industry and Its Workers* by Robert E. Parker  
ALISON DAVIS-BLAKE
- 1663 *The Contradictory College: The Conflicting Origins, Impacts, and Futures of the Community College* by Kevin J. Dougherty  
WALTER R. ALLEN
- 1665 *Neighbourhood Organizations and the Welfare State* by Shlomo Hasson and David Ley  
BRIAN K. GRAN
- 1668 **Book Notes**
- 1677 Acknowledgments to Referees
- 1681 Contents of Volume 100
- 1696 Book Reviewers for Volume 100

## IN THIS ISSUE

NADER SOHRABI is pursuing his doctorate in sociology at the University of Chicago. His research interests include the comparative study of political culture, revolution, and state formation.

SCOTT DAVIES is assistant professor of sociology at McMaster University in Hamilton, Ontario. His main interests are in the sociology of education and social stratification. His current research concentrates on school choice movements and other education reform groups in Canada.

CLIFF BROWN is a doctoral candidate in sociology at Emory University. His dissertation focuses on race and labor relations in the United States steel industry. He is also working on a related book with Terry Boswell.

TERRY BOSWELL is associate professor of sociology at Emory University. His current research includes a comparison of nine pivotal cases of union organizing and racial competition that span the period 1870 to 1950 in U.S. history. Other research projects include a reanalysis of the relationship between dependency and economic growth (with Bill Dixon) and a time-series analysis of revolutions in Europe since 1492.

MICHAEL HECHTER is university lecturer in sociology and fellow of New College, Oxford, and professor of sociology at the University of Arizona. He is currently writing a book on nationalism.

TIMUR KURAN is professor of economics and King Faisal Professor of Islamic Thought and Culture at the University of Southern California. His research focuses on the evolution of values and institutions, with current emphasis on ethnic conflict, cultural change, and the economic dimensions of religion. He has recently published several articles on the economic sources and accomplishments of Islamic fundamentalism and is the author of *Private Truths, Public Lies: The Social Consequences of Preference Falsification* (Cambridge, Mass.: Harvard University Press, 1995).

RANDALL COLLINS is professor of sociology at the University of California, Riverside. His books include *Theoretical Sociology* (1988), *Sociology of Marriage and Family: Gender, Love and Property* (4th ed., coauthored with Scott Coltrane, 1994), and *Four Sociological Traditions*. He is currently completing a project on the comparative sociology of philosophies.

CHARLES TILLY's most recent books are *Coercion, Capital, and European States* (rev. ed. 1992), *European Revolutions* (1993), *Cities and the Rise of States in Europe* (1994), and *Popular Contention in Great Britain* (1995). With Christopher Tilly, he is now writing *Capitalist Work and Labor Markets*.

EDGAR KISER is associate professor of sociology at the University of Washington. His research has focused on the determinants of war, the development and decline of voting institutions, and the evolution of tax systems. He is writing a book on the latter topic, tentatively entitled *Funding the State: Tax Administration in the Early Modern Era*.

JAMES S. COLEMAN is the University Professor of Sociology and Education at the University of Chicago. In addition to writing numerous books and articles, he is the founder and editor of the journal *Rationality and Society*.

ALEJANDRO PORTES is the John Dewey Professor of Sociology and International Relations at Johns Hopkins University. His most recent books are *City on the Edge: The Transformation of Miami* (coauthored with Alex Stepick; University of California Press, 1993) and *The Economic Sociology of Immigration* (Russell Sage, 1995). He is collaborating with Rubén G. Rumbaut on a longitudinal study of second-generation immigrants.

## Information for Contributors

**Editorial Procedures:** All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

**Manuscript Acceptance Policy:** While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

## Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures will be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract, on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or

money order for \$15.00, payable to The University of Chicago Press in U.S. currency or its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

**Format of References in Text:** All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is:

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

**Format of References:** List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 1/93)

# Historicizing Revolutions: Constitutional Revolutions in the Ottoman Empire, Iran, and Russia, 1905–1908<sup>1</sup>

Nader Sohrabi  
*University of Chicago*

The Young Turk Revolution of 1908, the Iranian Constitutional Revolution of 1906, and the Russian Revolution of 1905 unfolded under the impact of the constitutional revolutionary paradigm. This article places these revolutions in a common historical context, arguing that their shared ideology and their method of capturing state power differentiate them from revolutions of later periods. After establishing that these revolutions belonged to the same class of events, this article also explores the differences between the successful Ottoman and Iranian revolutions and the failed Russian revolution. The conclusion is that the Ottoman and Iranian legal assemblies, which were buttressed by extraparliamentary and extra-legal resources, were far more powerful and effective. Furthermore, the intrastate cleavages in the Ottoman Empire and Iran were essential for winning military support in the battle to restore the constitutional regimes in the aftermath of counterrevolutionary backlash.

## INTRODUCTION

### Revolutionary Paradigms, Stages, and Transitions

At the beginning of the 20th century, constitutionalism was the dominant revolutionary model. The actors in the Young Turk Revolution of 1908

<sup>1</sup> I am particularly indebted to Said Amir Arjomand, Şükrü Hanioglu, David Laitin, William Parish, Charles Ragin, William H. Sewell, Jr., and George Steinmetz for their incisive comments, suggestions, and encouragement. I also thank Andrew Abbott, Karen Barkey, Kathryn Babayan, Stephen Ellingson, Ivan Ermakoff, Benjamin Fortna, Roberto Franzosi, Douglas Glick, Mathew Payne, Bijan Sohrabi, Mehmet Ulucan, and Farzin Yazdanfar. Earlier versions of this paper were presented at the 1991 American Sociological Association's meetings in Cincinnati and the 1992 Social Science History Conference in Chicago. Dissertation research grants from Jerry Knoll/Wilder House and the MacArthur Foundation are gratefully acknowledged. Direct correspondence to Nader Sohrabi, Department of Sociology, University of Chicago, 1163 East Fifty-ninth Street, Chicago, Illinois 60637.

in the Ottoman Empire, the Iranian Constitutional Revolution of 1906, and the Russian Revolution of 1905 justified their claim to power by advocating constitutional systems of rule. Surprisingly, around the same time, actors in other revolutionary upheavals around the globe such as in China (1911) and Mexico (1910) made similar demands. The differing social structures of Russia, Iran, the Ottoman Empire, China, and Mexico strongly suggest that the ideology of constitutionalism could not have emanated from their social structures. Rather, the actors demanded constitutional systems of rule because they operated under the constraints of the "world time" that brought forth constitutionalism—the legacy of the French Revolution of 1789—as the dominant "revolutionary paradigm."<sup>2</sup>

In this article I argue that different paradigms, like Max Weber's "switchmen," set revolutions upon dissimilar trajectories (Weber 1946, p. 280). In each historical epoch, paradigms shape the conception of politics, orient the actors toward certain goals, and set limits upon the methods that may be utilized for realizing those goals. Thus, constitutional revolutions, influenced by the French paradigm, followed a trajectory quite different than the communist revolutions shaped by the Russian paradigm of 1917. By showing that communist and constitutional revolutions had entirely different dynamics and processes, the concept of revolutionary paradigm historicizes revolutions.

It is important here to clarify the relation between the French Revolution and the constitutional revolutionary paradigm and to define the latter against constitutionalism in general. The institutional history of constitutionalism is a complicated one, with origins long predating the French Revolution (see Downing 1988). Here I am not concerned with the institutional history of constitutionalism, nor am I particularly concerned with the history of the French Revolution. Instead, I am interested in constitutionalism as an ideological framework for revolutionary action. The French Revolution itself was ideologically informed by the history of British political institutions and the American Revolution, both of which were important for the formation of the French constitutional model. Yet rather than calling the paradigm British or American I call it French because it was the revolution in France that canonized the model, and the actors abroad referred more frequently to the revolution in France than to other upheavals with similar results. Furet expresses this idea concisely when he writes that until the French revolutionary model was cast aside by the 1917 Russian Revolution, France remained the "vanguard of history" (Furet 1981, p. 6).

<sup>2</sup> For arguments about "world time" see Fernand Braudel (1980).



For constitutional revolutionaries, assemblies and constitutions were models without history. The actors considered the French model a “blue-print” and did not conceive it to be a historical product of particular and unique circumstances. They did not approach French history as historians but as politicians, manipulating and molding the French Revolution to their own ends. Reinhard Bendix (1984, pp. 114–16) pointed to this idea when he wrote that the French Revolution is one of the best examples of a timeless event with “demonstration effect.” Bendix argued that, regardless of the prehistory of the French Revolution and the unique combination of factors in France that led to its outbreak, “once the French Revolution had occurred, other countries could not and did not recapitulate that prehistory; they reacted to the revolution itself instead” (p. 116).<sup>3</sup> Similar to Furet, he held that after the Russian Revolution of 1917, the Russian model superseded the French one by becoming the new “reference society” (p. 116). Arjomand’s study of constitutionalism also points to the timelessness and autonomy of models and reaffirms this claim when he notes that “the institutional structures and normative patterns generated in the formative experience of one nation become blueprints autonomous of the particular circumstances of their birth, and acquire fixity and rigidity” (1992, p. 39; see also p. 73).

Constitutional and communist revolutions differ; constitutionalists did not demand a sudden and complete overthrow of the institutions of the old regime but instead called for the creation of an elective representative body through which they attempted to indirectly dominate the state. They desired to gradually reform the political structures of rule—the state and social institutions—while also changing the locus of sovereignty and state power. They attempted to create legal-rational administrations and to put in operation a written constitution that outlined new authority structures and a separation of judicial, executive, and legislative powers. One may, perhaps less satisfactorily, label these events as parliamentary revolutions. In short, the constitutional ideological structures, rather than advocating a sudden and total takeover of state institutions, defined the goal of the movement to be creation of a strong representative, legislative assembly that dominated the executive. It is this distinct organizational and incremental aspect of the constitutional revolutions that sets them apart from the initially more violent communist or socialist revolutions and gives them an altogether different dynamic.

Though some have acknowledged the significant differences and the historical supersession of the Russian with the French model, the ways that revolutionary models or paradigms organize conflicts have not been

<sup>3</sup> On the effect of the French Revolution see also Arendt (1977, p. 55).

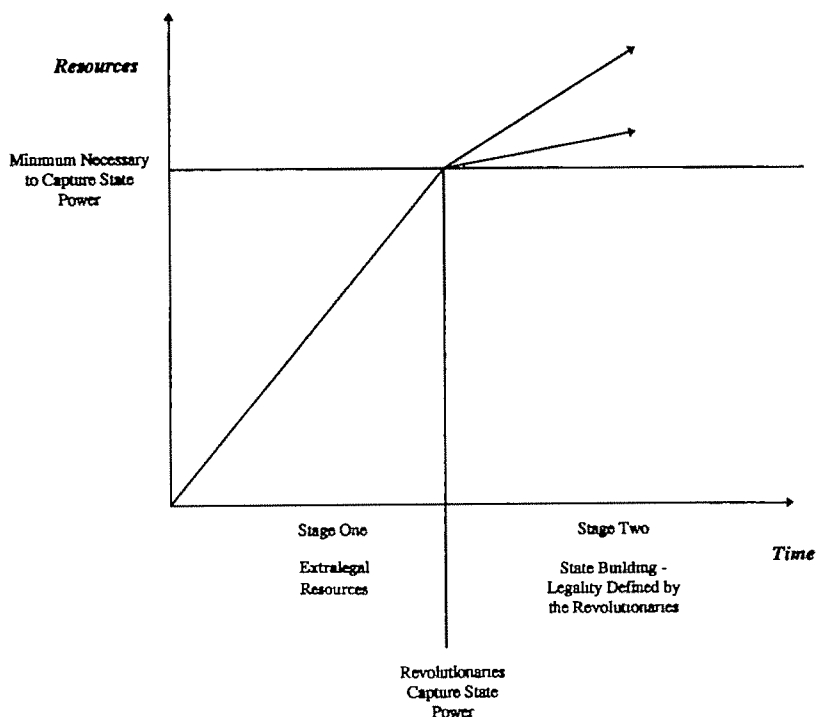


FIG. 1.—Communist/socialist revolutionary processes

demonstrated in an empirical and systematic manner. This study's empirical demonstration of this paradigm effect is aimed at filling the void. The ideal-type of a socialist or communist revolution is demonstrated in figure 1. Broadly speaking, this type of revolution is divided into two stages and one transition period. In the first stage, the contenders create power blocs to wage a fight for state capture by mobilizing extralegal resources. In the successful case, transition occurs when the revolutionaries muster enough resources to capture state power and dismantle the old regime, at which point the battle concludes and the revolutionaries begin to introduce radical reforms and impose their own definition of legality. Radical reforms are introduced only in the second stage, when the revolutionaries, at least formally, are in state command and the old regimes have been deposed. Having full control of the coercive state organs, and imposing their own definition of legality, the communist or socialist revolutionaries are much less vulnerable to a counterrevolutionary backlash.

In contrast, as figure 2 suggests, constitutional revolutions are divided into four distinct stages and three transition points. In the first period, the

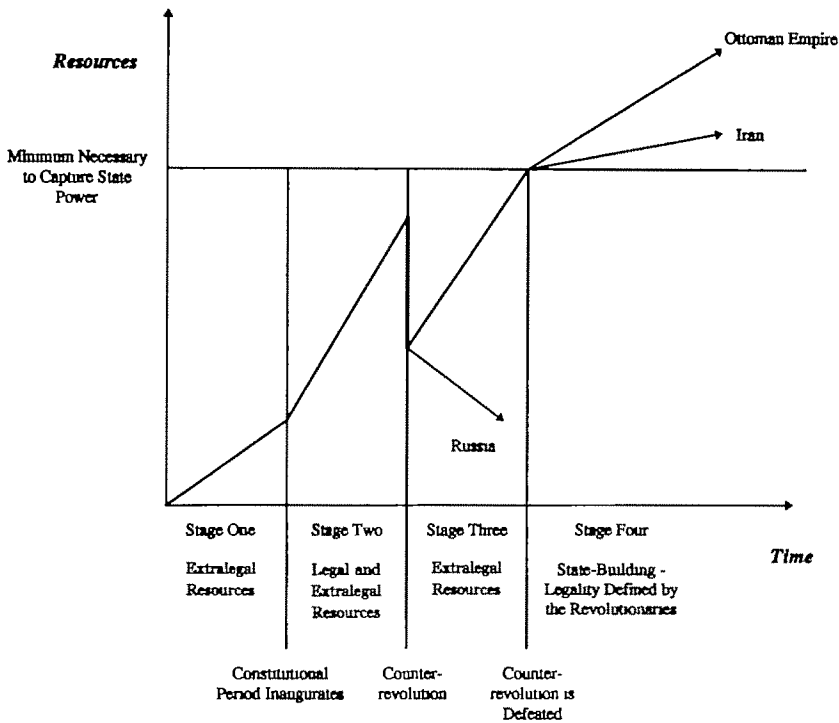


FIG. 2.—Constitutional revolutionary processes

constitutional revolutionaries create power blocs by mobilizing extralegal resources, but rather than aiming for complete overthrow of the state, they demand creation of legislative assemblies. Once the threatened old regimes agree to this demand, a period of legal activity ensues. What is distinct is that the compromise between the old regimes and the assemblies is reached long before the constitutionalists gather enough strength to take over the state. But, once the compromise is forged, the second stage provides ample opportunity for further consolidating more resources for the power blocs; not only does the cost of mobilizing previously extralegal resources decrease, but constitutionalists are also given a greater chance to increase their support from within the state ranks and to construct quasi-governmental institutions in support of legally sanctioned assemblies. The assemblies, once established, begin to implement or demand radical reforms that gravely disturb the old regimes and a wide range of social groups, all at a time when the constitutionalists lack total state control. The old regimes, which still hold the upper hand in state control, in cooperation with the adversely affected social groups, respond by un-

leashing a counterrevolution that forces the assemblies out of power and creates an atmosphere of illegality resembling the preconstitutional period.

The above contrasts demonstrate in broad terms the major differences of the constitutional and socialist revolutions. My aim in this article, however, is not only to establish that constitutional revolutions have similar processes distinct from socialist revolutions, but also to compare instances of two successful constitutional revolutions with a failed instance of constitutional revolution to causally account for the observed difference. Contrasting the two successful Ottoman and Iranian revolutions with the failed attempt in Russia, I conclude that, ironically, the degree of legislative success of the legal assemblies and their ability to implement their programs depended upon the extraparliamentary, and in many instances extralegal, support they received from within the state or from the quasi-governmental institutions created during the second stage. Furthermore, the Ottoman and Iranian constitutionalists recovered from the counterrevolutionary backlash because they enjoyed the support of quasi-governmental institutions, which were strengthened during the period of legal activity, and even more critically, because they enjoyed the support of a sector of the armed forces. With this assistance the constitutionalists advanced to the fourth stage by capturing state power, deposing monarchs, and restoring the constitutional regimes. The Russians who commanded neither of these resources were left with a constitution only in name and an assembly that, after various old regime-sponsored legal modifications, was transformed to an utterly powerless institution.

#### Ideologies, Revolutionary Processes, and Outcomes

Theories of revolution make a useful distinction between processes and long-range outcomes, yet these theories differ in the importance they attach to the impact of ideologies upon processes or outcomes. Some argue against the impact of ideologies. For example, insisting on a "structuralist" approach, Skocpol dismisses theories that give critical importance to ideologies by pointing to an exaggerated emphasis on agency and voluntarism (see Skocpol 1979, pp. 14–18, 164–71; 1976, pp. 209–10). Others such as Sewell emphasize the impact of ideologies on both outcomes and processes. For Sewell, fundamental divergences in the outcomes of "bourgeois" and "socialist" revolutions are explained by differences in the actors' ideologies (Sewell 1985, p. 59; for a rejoinder, see Skocpol [1985]). In agreement with Sewell, and with an added emphasis on ideological transformation throughout the revolutionary process, Goldstone emphasizes the effect of ideologies on postrevolutionary recon-

struction (Goldstone 1991a, 1991b). The debate on processes has for the most part been concerned with ideologies' causal impact on the downfall of old regimes. Sewell (1985), for example, argues that the ideological contradictions under the old regime in France contributed immensely to its downfall. Similarly, Kimmel (1990, pp. 185–87) is concerned with how ideologies have a direct causal impact on the downfall of old regimes.

In this article, I am not concerned with long-range outcomes but with the impact of ideologies upon processes. Like Sewell, I side against those theories that discount the impact of ideologies, but my approach differs from Sewell's in that I do not discuss the causal impact of ideologies upon the collapse of old regimes. Instead, I am concerned with how ideologies pattern the stages that revolutions pass through from their inception to the final capture of state power by contenders. I argue that ideologies have a definite and visible impact on processes, and I maintain that if revolutionary actors operate under the influence of the same revolutionary paradigm, processes become divided into stages that are analogous and also quite distinct from the stages under other, alternate revolutionary paradigms. Investigation of causality in the comparative study of revolutions may begin only after various stages of revolutions have been delineated, stages that are largely shaped by revolutionary ideologies.

Although Skocpol's "nonvoluntarist" account is also sensitive to the concept of "world time" and demonstration effect, her empirical discussion reduces their significance to institutional borrowing, such as the adoption of Leninist party structure during the Chinese revolution (Skocpol 1979, pp. 23–24). In contrast, the concept of revolutionary paradigm not only highlights the importance of institutional borrowing from one revolution to another but has a significance that goes far beyond mere institutional borrowing.<sup>4</sup> It shows that revolutions that operate under different paradigms are dissimilar events.

### Methodology

Two instances of success and one instance of failure make Mill's method of agreement and indirect method of difference an appropriate methodology for this discussion (see Skocpol and Somers 1980; Ragin 1987, pp. 34–52). Yet, as Ragin has argued, within the framework of Mill's indirect method of difference, it is logically very difficult to define proper negative instances of phenomena such as social revolutions (Ragin 1987, pp. 41–42). For a meaningful comparison, what one should be able to demonstrate here, therefore, is that the events in the Ottoman Empire, Iran,

<sup>4</sup> For arguments that give greater weight to revolutionary models, in addition to the above, see Sewell (in press) and Hermasi (1976).

and Russia constituted similar instances of constitutional revolutions. I demonstrate this by emphasizing their largely similar initial processes and the similar dynamics of unfolding. Once it is demonstrated that they all shared similar stages, but that the Russian revolutionaries failed to advance to the final stage and capture state power, then we can claim, with comfort, that Russia constituted a negative instance of constitutional revolution. The challenge is therefore twofold: to argue that all three revolutions were similar enough to be considered one class of phenomenon, namely constitutional revolutions, and then to show that Russia represents the failed instance while Iran and the Ottoman Empire exemplify the successful cases. This strategy allows comparison in order to causally account for similarities and differences.

To demonstrate their similarities, I adopt a formal narrative approach to argue that in all three events the revolutionary dynamics were analogous: they all passed through parallel stages, they experienced comparable turning points, and the causal forces propelling the transition from one stage to another were largely similar (for more on the narrative approach, see Sewell [in press] and Abbott [1983, 1990, 1991]). Emphasis on the processes, stages, and narratives in the study of revolutions has been traditionally associated with the natural historians (see Edwards [1927] 1970; Brinton [1938] 1952; Pettie 1938). Despite their sophisticated methodological assumptions (see Abbott 1983, 1992), these "stage" theories fail to provide analytical reasoning as to why revolutions are divided into various stages, what marks the boundaries of various stages, and more important, why and how transition from one stage to the next takes place. Instead, they confine themselves to a descriptive analysis of each stage (for a review, see Goldstone [1982, pp. 189–92]). In contrast to the natural historians, I am centrally concerned with the logic behind the division into various stages, the boundaries that mark the stages, and the comparative analysis of causal forces that propel the transition from one stage to another. Thus, while my approach stresses narrative, it is distinct in its explicit causal-analytic emphasis.

In the following, I present a selective historical narrative account of the revolutionary politics under stages 1, 2, and 3. But before doing so, I highlight the revolutionary context by describing the ambitious prerevolutionary undertakings of (autonomous) states within the military, administrative, and social structures.

#### PREREVOLUTIONARY REFORMS OF THE 18TH AND 19TH CENTURIES

The manner in which prerevolutionary reforms in the 18th and 19th centuries affected state and social structures helps explain why large

sectors of the Ottoman army and bureaucracy, and a sector of the Iranian army, sided with the opposition against the old regimes. It also sheds light on why the Russian bureaucracy and army remained loyal to the Tsar. Here I argue that preexisting structural cleavages within the military and state institutions were good indicators of whether their members sided with the opposition.

In this analysis, the prerevolutionary state reforms are considered to be a failure in Iran, only a partial success in the Ottoman Empire, and a complete success in Russia. It is interesting that armies were the locus of the first waves of modernization in all three societies. Defeat in war at the hands of superior armies was a potent impetus for the reformers to modernize their armed forces and, gradually, their civil bureaucracies. It should also be noted that the degree of the general success of prerevolutionary modernizing reforms did not have a simple linear impact on the vulnerability of these societies to revolution. In other words, more success at reforms did not simply translate into less vulnerability to revolution. Furthermore, to assess the impact of prerevolutionary reforms on revolution their effect on state and social structures must be evaluated individually.

By the time of the state financial crisis and foreign political pressures of the early 20th century, the Ottoman reformers had managed to transform the civil and military administrations, but only partially. The era of modern reforms began after the Ottoman defeat in the Russian War of 1787–92 and continued until 1908. The reforms of the early 19th century, particularly those under Mahmud II (1808–39), paved the way for the creation of a weak civil society. By first suppressing the Janissaries and later their close allies, the popular *Bektashi* religious order, Mahmud II initiated a process of weakening various popular institutions with ties to guilds, artisans, and other social sectors, a process that continued to the end of the reform era. Mahmud II also managed to quash the provincial notables and, with them, the provincial autonomy of almost all regions with the exception of Egypt. Weakened networks of social actors and the destruction of some of their key institutions left the Ottoman civil society vulnerable to state encroachment and made the state much less susceptible to a popular mass uprising from below (see Lewis 1961, pp. 78–83; Shaw and Shaw 1977, pp. 19–24, 41–45; Brown [1867] 1927, pp. 163–64; Birge [1937] 1965, pp. 16, 77; Heyd 1961, pp. 64–69; Kissling 1954).

Most significantly, the later reforms of 1839–1908 managed to create major modernized sectors within the Ottoman military and bureaucracy, sectors that began to operate on the basis of legal/rational rules of conduct. Yet, at the beginning of the 20th century, large areas of the patrimonial structure remained intact and the Ottoman state functioned under

two differing administrative regulations and rules of promotion.<sup>5</sup> The severe structural division within the Ottoman army and bureaucracy nurtured strong grievances among the rationalized sectors of these institutions, thus prompting them to take revolutionary action against members of the traditional/patrimonial bureaucracy who were deemed to have blocked the mobility of the younger, military and civil bureaucrats with modern educations. The formation of a "critical mass" of civil and military officers who desired a rational administration, together with the presence of a "reference group" of patrimonial officials who operated alongside these officials, created strong grievance-generating mechanisms within the Ottoman administration, turning modernized sectors into members of opposition. With the financial and political weakening of the state in the early 20th century, these sectors became revolutionary.

In Russia, on the other hand, similar structural divisions did not exist within the army or the bureaucracy. The reforms that Peter the Great (1682–1725) had initiated in the aftermath of Russia's defeat by Sweden were succeeded by a new series of reforms under Alexander II following the Crimean War. The most significant of Peter the Great's reforms was the introduction of tables of ranks, which despite its slow beginning, gradually affected the entire administration. At the beginning of the 20th century, when Russia entered a period of political instability marked by state financial crisis and foreign threat, the reforms had managed to turn the state administration into a uniform and modern institution. Although the autocracy was anachronistic by Western European standards of the time, roughly similar legal rational rules prevalent in other European administrations predominated in the Russian bureaucracy and the army. At the end of the 18th century, the importance of family connections, training on the job, and mixed military and civilian careers had given way to objective criteria such as formal education.<sup>6</sup> Thus, unlike the Ottoman military and civil bureaucracy, the corresponding institutions in Russia were not riddled with structural divisions between modernized and patrimonial officials who operated on the basis of conflicting rules of conduct.<sup>7</sup>

The reforms in Iran, like those in the Ottoman Empire and Russia, had their roots in martial defeat and first concerned the military. More-

<sup>5</sup> For history of administrative reforms within the Ottoman state see Findley (1980, 1989), Lewis (1961), Shaw and Shaw (1977), Berkes (1964), and Davison (1963).

<sup>6</sup> For an extended treatment of the Russian bureaucracy and its rules of conduct see Pintner and Rowney (1980), Fainsod (1963), Yaney (1973), and Raeff (1984).

<sup>7</sup> Even the most skeptical observers of Russian administration do not attribute the divisions within the autocratic, and what is portrayed as a nonrational bureaucracy, to the conflict between "modernized" and "traditional" bureaucrats (see Verner 1990, pp. 44–69, 167–74).



over, the reforms in Iran were initiated under the direct influence of reforms in Russia and especially the Ottoman Empire, but even compared to the latter they were far less successful. Two attempts to organize a modern standing army after two Russian defeats in 1813 and 1827 bore only limited success and were disbanded shortly after. The administrative reforms did not fare any better. Although European forms of administration were introduced, they failed to fundamentally transform the state's decentralized structure or to change the essence of the central state's traditional structure into a modern state with distinct ministries and a well-defined division of duties. Thus unlike the Ottoman bureaucracy, the Iranian bureaucracy was not rife with divisions of rational and patrimonial offices but remained largely a uniformly patrimonial entity (Arjomand 1988, pp. 24–26). As a result, in contrast to the Ottoman state, where the modernized staff constituted a weighty discontented force, the few modernized officials that found their way into the state in Iran did not constitute a critical mass with similar grievances against the patrimonial officials. The lack of clear and fundamental divisions within the state meant the absence of grievances on the part of large numbers of officials within it; this in turn translated to the loyalty of officeholders to the monarchy at the time of serious political crisis. Thus, ironically, the undivided nature of the Iranian patrimonial bureaucracy made it the “functional equivalent” of the rational, but similarly undivided, Russian bureaucracy.

On the other hand, if the failure of reforms left the administrative structure at the center intact, the same failure left the Iranian army a structurally divided entity. The reforms failed to destroy the tribal nomadic contingents and replace them with a strong standing army at the center, a situation that remained true even under the best of reformers in the 1870s.<sup>8</sup> Between 1880 and 1907, other than drastic deterioration of its material and training conditions, the structural characteristics of the army did not undergo any major transformation and it continued to be composed of “(i) the regular infantry; (ii) the tribal levies, chiefly mounted; and (iii) the artillery, or rather men enlisted as gunners.” The only major change was the addition of the Russian-trained and controlled Cossack Brigade that, although better organized, trained, and equipped than the rest of the Persian army, was a small group (1,500 in 1899) suitable only for guarding and policing Tehran (Ra'iss Tousi 1988,

<sup>8</sup> As Bakhsh (1978, pp. 98–100, quote at p. 100) has noted, “The army continued to be composed, as it had traditionally been under the Qajars, of a permanent royal guard, an irregular cavalry based on tribal levies, an irregular infantry militia raised and supported locally by each district and a semi-regular army of infantry, cavalry and artillery which constituted the bulk of the defense forces.”

p. 209). Thus, the loose alliance of small divisions of the standing army in the capital and provinces with the semiautonomous tribal nomadic contingents in the periphery remained a feature of the Iranian army that lasted into the revolutionary era of the early 20th century. Such loose alliances, at any point and for a variety of reasons, could be severed and overturned. This was especially true during the Qajar period, as the state attempted to break the military power of tribal groups and replace them with a standing army (Arjomand 1988, p. 23).<sup>9</sup> The Qajar reformers also left the institutions of civil society, most significantly that of religion, intact. This gave the social actors in Iran a powerful channel of protest, a significant factor that was absent in the Ottoman Empire, for the reformers there had succeeded in destroying vital institutions within civil society.<sup>10</sup>

It is important to describe in broad terms the history of reforms prior to the revolutions; it highlights the structural cleavages within the administrative and military institutions, divisions that became salient during the revolution. I now turn to an analytical history of various stages of these revolutions. Contrary to expectation, I will not begin this analysis with the discussion of the uprisings that forced the old regimes to make concessions. These uprisings involved many actors with disparate grievances and interests. Recounting them in detail, though important for some purposes, diverts from the essential focus of this article. Instead, I concentrate on the transition point between stages 1 and 2 to answer the following: First, why did actors with diverging interests agree to the same principle of constitutionalism and, second, why did the old regimes enter a compromise at such an early time, long before confronting the threat of a total takeover of their administrations? The answer, I suggest in the following, lies in the ambiguities of the concept of constitutionalism.

#### STAGES 1-2: IDEOLOGICAL AMBIGUITY INTENDED AND UNINTENDED

Constitutional ideologies had ambiguous and unambiguous features that were responsible for permitting the transition from stage 1 to stage 2 and

<sup>9</sup> Note that Arjomand and Ra'iss Tousi present widely varying pictures of the military power of the tribes at this period. Ra'iss Tousi holds that between 1880 and 1907 the tribes were very well equipped and owned over twice as many arms as the Persian infantry, which "clearly reveals the superiority of tribes over the infantry" (1988, p. 215). Arjomand, on the other hand, holds that the tribal power was effectively broken in this period, even though it was not replaced with a standing army at the center.

<sup>10</sup> For reforms in Iran see Bakhsh (1978, 1983), Arjomand (1988), Adamiyat (1969, 1972), Nashat (1982), and Floor (1983).

opening up the path of action that gave constitutional revolutions their characteristic trajectory. What was clearly unambiguous about the constitutionalists' goal was that, unlike socialist or communist revolutionaries of later eras, they did not strive for a sudden takeover and complete overthrow of the traditional structures of rule. Instead, they demanded the establishment of assemblies with power over the executive branch, the transference of the locus of sovereignty, the reform of the state structures, and the implementation of radical political and social programs. These measures were sweeping and drastic and were resisted by the old regimes. Yet, once confronted with serious uprisings demanding constitutional systems of rule, the old regimes did agree to them, and did so surprisingly early.

The old regimes' strategy, however, was to agree to a set of ambiguous principles that promised the establishment of institutions resembling, but not entirely equivalent to, parliamentary systems. By entering into a compromise over ambiguous principles, the old regimes hoped to end the various forms of protest, such as the military uprising in the Ottoman Empire, the large-scale and pervasive sanctuaries in Iran, and the massive waves of strikes in Russia. The promise to allow some form of an assembly allowed them to buy more time at a critical juncture and to abate the tide of revolution. In fact these compromises accomplished both goals, but old regimes were then forced to spend great effort during the second revolutionary period to regain what they had lost as a result of their ambiguous promises. On the other hand, during stage 2 the opposition attempted to impose its own definition on the terms of the compromise and to offer its own version of the agreement.

Such ambiguity is evinced in the issued decrees that supposedly commenced the constitutional systems of rule. For example, in the Ottoman Empire, in reaction to the soldiers' uprising in the Western provinces, the sultan's decree issued on July 23, 1908, promised the establishment of the Chamber of Deputies in accordance with the constitution of 1876.<sup>11</sup> Yet the decree left many issues unresolved as the constitution of 1876 (as will become clear below) had left many issues with regard to the power of the assembly ambiguous. Furthermore, the revolutionaries were not provided with any convincing guarantees that the constitution would be put into effect. The Ottoman state, when confronted with a constitutional movement in 1876, had agreed to a compromise by granting a constitution and a chamber of deputies. However, after the accession of Sultan Abdülhamid II and the renewed restoration of imperial authority in the

<sup>11</sup> *Düstur*2, I, July 23, 1908 (24 Cemaziyelahr 1326), pp. 1–2.

same year, the new sultan reneged on the latter promise (Mardin 1962, pp. 56–78). In addition, despite violating the central tenet of constitutionalism, the sultan never annulled the constitution, consistently conveying that the Ottoman state was indeed a constitutional state (beginning in 1877, *Salname*, the official yearbook of the state, had unfailingly published the text of the 1876 constitution). Without doubt, by agreeing to an early compromise in the summer of 1908, the Ottoman sultan was intending to imitate the delaying tactics used in 1876.

In Iran, the shah's decree that came to be interpreted as the order for the commencement of the constitutional system of rule was perhaps the most ambiguous document of its kind. The wording of the decree dated August 5, 1906, likened the National Consultative Assembly to an advisory panel of reform rather than a legislative assembly. Furthermore, it failed to mention the word "constitution."<sup>12</sup>

Similarly in Russia, the decree issued by the Tsar Nicholas II on October 17, 1905, known as the October Manifesto, intentionally avoided using the word "constitution" (Ascher 1988, pp. 228–29, 231; Healy 1976, pp. 108, 275–76). While not as ambiguous as the shah's decree, it was notoriously vague on the governing powers of the assembly, its legislative duties, and its representative character. Even though the announcement of the October Manifesto was followed by large-scale outpourings celebrating the grant of a constitutional system of rule, as it was done in the Ottoman Empire and in Iran, many leaders of the Russian opposition felt that the manifesto left many critical issues unreasonably ambiguous; it had taken away with one hand what it had given with the other. The Fundamental Laws, issued on April 23, 1906, weary of the European word "constitution," also rejected this vocabulary and, like the manifesto, remained vague on many critical issues (Ascher 1988, pp. 231–32, 242; 1992, pp. 12, 16, 60–61, 63–71, 190–91; Healy 1976, pp. 15–22, 58–62, 81–82, 106, 108, 264–65; Harcave 1964, pp. 130, 161–62, 195–96, 199–202; Hosking 1973, pp. 10–13, 54–55.)

Furthermore, in all instances, the old regimes made it known that the constitutional systems were granted out of the free will of the monarch.<sup>13</sup> This line of argument, of course, had ominous implications for the oppo-

<sup>12</sup> The decree was dated August 5, 1906 (14 Jumada II 1324) to make it correspond to the shah's birthday. It was in fact issued four days later (Nazim al-Islam Kirmani 1983, 1:561–64).

<sup>13</sup> For the Ottoman Empire see *Fikir Hareketleri*, no. 100, September 21, 1935, p. 343; *İkdam*, no. 5320, March 16, 1909 (23 Safer 1327), p. 1; Ahmad (1969, p. 13). For Iran see Nazim al-Islam Kirmani (1983, 1:561–64, 628–30); *Maylis* September 27, 1906 (8 Sha'ban 1324), p. 8. For Russia see Ascher (1992, pp. 12, 60–61).

sition: if the constitutions were granted out of the monarchs' pure benevolence, they may just as easily be removed.

In part the old regimes were able to issue ambiguous decrees because members of the opposition had themselves built coalitions based on ambiguous principles. All were united on the basis of their opposition to the old regimes, but their interpretation of the political system they desired and their degree of antagonism toward the institutions of the old regime varied widely. Thus, the "particular" grievances of a diverse array of actors toward the old regime, actors who had varying and often conflicting interests, were expressed in terms of the "general" demand for constitutional systems. The ambiguities that remained within such a general demand were essential for building broad-based coalitions among groups with incompatible interests and varied interpretations of the new political system.<sup>14</sup> The range of revolutionary actors in Iran was diverse: the guilds and tradesmen, merchants, clergy, landlords, some sectors of the peasantry, and a small section of the statesmen of the old regime. In the Ottoman Empire oppositionists were professionals, university students, various nationalist minority groups, some sectors of the peasantry, and, most significantly, large sectors of the modernized military and civil bureaucrats. In Russia, the opposition was composed of the working class, peasants, professionals, university students, and the landowning gentry.<sup>15</sup>

While in all settings constitutionalism was introduced and understood by many actors as a European political system and ideology, it was also "translated" and adapted to existing local traditions or presented as the solution to many particular problems in each setting. In the Ottoman Empire and in Iran, for example, where economic, social, and institutional "backwardness" were major concerns, the ideology of constitutionalism was introduced with strong overtones of progress that among other things meant economic advancement, a modern state, and a legal-rational order. Furthermore, as an example of adaptation of constitutionalism to local traditions, in Iran the guild members and the clergy equated the establishment of the assembly with the rejuvenation of the "House of Justice."<sup>16</sup> In the Ottoman Empire, where ethnic strife and nationalist movements were some of the most pressing issues, constitutionalism was

<sup>14</sup> Gene Burns (1991) has advanced a similar argument for the 1979 revolution in Iran. For a similar argument about progression of ideologies from "particular" to "general" prior to revolutionary outbreaks see Goldstone (1982, p. 203).

<sup>15</sup> The broad-based coalitions and contradictory interests of actors in 1905 Russia have been acknowledged by many scholars (see, e.g., Ascher 1988, p. 244, Harcave 1964, p. 12; Hosking 1973, p. 3; Verner 1990, p. 3).

<sup>16</sup> Nazim al-Islam Kirmani (1983, 1:358–66); Adamiyat (1976).

presented as an ideological companion to, or even a substitute for, the failing Ottomanism—an ideology that professed peace and harmony among the subject populations of the empire.<sup>17</sup> In Iran and the Ottoman Empire as in Russia, the peasants interpreted “liberty,” a central slogan of the constitutional movement, to mean freedom from many traditional obligations, which in some cases even meant paying taxes (Ascher 1988, p. 233).

Thus, calls for constitutional systems of rule, even when the actors insisted on constitutional monarchies, left the field open to more radical suggestions. By granting assemblies, the monarchs opened themselves to the danger that the assemblies might be turned into fully legislative and representative bodies, and that such powers could then be used to depose the reigning monarch or, worse yet, to abolish the monarchical systems in general. On the other hand, the ambiguous compromise left the old regimes sufficient room to deny the new assemblies substantive powers. After the compromise, the fight opened itself to extreme possibilities, and debates raged around whether the assemblies should be confined to purely consultative bodies or full-fledged parliaments. Stage 2 of the revolution was devoted to working out these ambiguities. Revolutions were formally announced after relatively bloodless confrontations, but the real battle was yet to be fought—a struggle that was to reach its fiercest moments in the second and third revolutionary stages.

#### THE LEGAL PHASE: RISE IN RESOURCES AND ESTABLISHMENT OF THE ASSEMBLIES

The existence of a legal period of revolutionary activity distinguishes the constitutional from the socialist (or communist) revolutionary processes. For Charles Tilly, the “revolutionary situation” begins with the emergence of two or more centers of power (multiple sovereignty) and it ends when only one center of power gains control over the government (Tilly 1978, pp. 191–208). Tilly, however, fails to make a qualitative distinction between power blocs during constitutional and socialist revolutions. The power blocs that arose during constitutional revolutions became legally recognized by the old regimes and were regarded as legitimate not only by a significant portion of the population (Tilly 1978, pp. 191, 200) but by the old regimes as well (see fig. 2 above). Power blocs that emerged

<sup>17</sup> See the resolution of the Second Young Turk Congress in Kuran (1945, pp. 238–43, esp. 239–240). For similar views in one of the earliest resolutions and that of the First Young Turk Congress see Tunaya (1952, pp. 117–18, 123–27). See also *Şura-yi Ümmet*, no. 18, December 16, 1902 (15 Ramazan 1320), p. 4.

during socialist or communist revolutions certainly did not enjoy legal status during revolutionary situations.

With such compromise, the revolution entered its second legal phase, during which assemblies were established. The power of legal assemblies in their infancy, ironically, rested upon the extraparliamentary resources they could muster, which in many cases included extralegal resources as well. After the compromise, the old regimes were not simply willing to abide by the assemblies' legal orders; their respect for these institutions rested instead upon their balance of power with the assemblies. The state power, allowing for variations in each case, was to a large degree still in the hands of the incumbent governments, and they each obeyed the assemblies' orders only when they felt sufficiently threatened.

Entering the second stage affected the challengers' resources in a strikingly positive way. In this stage, many previously illegal activities became sanctioned by law, and the constitutionalists used this period to augment the extraparliamentary resources of the assemblies. Yet, because of the particular structural characteristics of Russia, Iran, and the Ottoman Empire, not all revolutionaries benefited equally from the opportunities of stage 2. In one extreme stood the Young Turks, revolutionaries who bolstered their power to the greatest extent when they used their newly won rights to establish a firmer foothold within the state and to gain a significant share of state power. In the other extreme stood the Russian revolutionaries, whose institution-building activities at this stage, though impressive, were far from sufficient to seriously challenge the state. The Iranians could not penetrate the state ranks, yet they had great success in building quasi-governmental institutions and organizing an extensive militia.

In the Ottoman Empire, because of internal state support for the Young Turks, the Chamber of Deputies' threat was real enough and was obeyed to a greater extent than those of the Iranian National Assembly or the Russian Duma. In Iran, even though internal state support for the constitutionalists was meager, the quasi-governmental forces were intimidating enough for the weak government to take the assembly seriously, to act upon some of its demands, and to negotiate with it on various issues. In contrast to both, the Russian Duma was established after the government imposed greater restrictions on the opposition's legal activities and crushed the quasi-governmental forces that supported the Duma. As will be discussed below, these differences proved to be major reasons for the comparative weakness of the Russian Duma. In Russia, because the Duma enjoyed little extralegal support, the government could afford to ignore its demands almost entirely and tamper with its institutional structure to reduce its legal power.

The Powerful Ottoman Chamber

After the Ottoman sultan granted a constitutional system of rule on July 23, 1908, the center of subversive activities shifted from the western provinces and European capitals to Istanbul. The opposition no longer had to fight the government clandestinely but could do so openly and legally using their newly won rights. This task was facilitated by the unleashed opposition press that, already in place in foreign capitals, now began to legally operate in Istanbul.

The powerful Ottoman Chamber of Deputies was also established in this period. To understand the chamber's sources of strength, we have to look at the composition of the Young Turk opposition. The Young Turks were composed of a variety of groups opposed to the sultan, the most powerful of which operated under the Committee of Union and Progress (hereafter CUP). The constitutionalist CUP was composed predominantly of large numbers of young military officers and civil administrative employees with modern education who opposed the traditional/patrimonial army and bureaucratic staff. The Ottoman Chamber of Deputies drew the majority of its members from the CUP, and this situation allowed the chamber to tap into the strength of an organization that had a firm foothold within the state. The chamber increased this support through a legal campaign of administrative reorganization and purges within the army and the civil administration, a campaign that proportionally increased and advanced the CUP ranks within these institutions, giving them the upper hand. A further addition to the chamber's strength was when the CUP, using the reduced mobilization cost of stage 2, infiltrated more deeply in the provincial administration and expanded its semiclandestine clubs.

Yet, even before the opening of the Chamber of Deputies, the CUP proved its presence within the state and its influence upon the old regime in many different ways. On August 6, 1908, two weeks after the proclamation of the constitutional system, the CUP forced the resignation of the "old style" grand vizier, Said Pasha, and pressured the government to appoint Kamil Pasha, who was thought to be in conflict with the palace and in agreement with the CUP's political programs.<sup>18</sup> When the CUP found itself in strong disagreement with Kamil Pasha as

<sup>18</sup> For disagreements with Said Pasha's government see *Fikir Hareketleri*, no. 76, April 4, 1935, "Meşrutiyet Hatıraları," pp. 374–75 (since all articles bear the same title, future references will be to the publication only); *Tanin*, no. 3, August 4, 1908 (6 Recep 1326) pp. 1–2; no. 5, August 6, 1908 (8 Recep 1326), p. 3, and pp. 3–4. These latter two *Tanin* articles also contain initial statements of support for Kamil Pasha (esp. p. 3). See also *Tanin*, no. 15, August 16, 1908 (18 Recep 1326), p. 1. For statements of support that appeared in *Sabah*, see Ahmad (1969, p. 21).



well,<sup>19</sup> by means of its palace associates, the CUP applied direct pressure on the sultan to dismiss him. The sultan succumbed to the pressure. The only matter that kept Kamil Pasha, the influential Ottoman grand vizier, from being dismissed immediately was that the CUP awaited the opening of the Chamber of Deputies to give an appearance of legality to its illicit activities. After the opening of the chamber on December 17, 1908, the grand vizier's minor request to postpone for four days a response to the chambers' interpellation served as a pretext to dismiss him by a broad voting margin on February 13, 1909. This action was a clear show of force, for the CUP had proved its influence even without the assistance of the chamber.<sup>20</sup> Another show of strength came in the early days of the constitutional government when the CUP ordered the anchoring of a navy ship in front of the palace with its guns trained on the sultan's residence; no one had the authority to order the ship's removal.<sup>21</sup>

The Chamber of Deputies began its operations on December 17, 1908, and one of the earliest pieces of legislation it approved was the law on purges. The legal, CUP-directed purging campaigns within civil officialdom and the army enabled the CUP to penetrate deeper into the state and to create an even more secure foothold. After purging the officials and army officers who had risen through ranks by means of connection to the patrimonial household of the sultan or a grandee, the CUP opened room for its supporters within the state and the army. By advancing officials and officers with modern training, the CUP further consolidated its position within the state.

After the granting of the constitution, the CUP expanded its prerevolutionary network to establish an organizational structure that resembled and competed with that of the state. Taking advantage of its legal status, the CUP organized more extensively around the empire and expanded the semiclandestine activities of its clubs and party branches and took a more active role in provincial administration. The CUP operated not

<sup>19</sup> For the CUP's later disenchanted view of Kamil Pasha see *Tanin*, no. 128, December 8, 1908 (14 Zilkade 1326), pp. 1–2; no. 129, December 9, 1908 (15 Zilkade 1326), pp. 1–2; no. 145, December 25, 1908 (1 Zilhicce 1326), p. 1; no. 157, January 8, 1909 (15 Zilhicce 1326), p. 1; no. 160, January 11, 1909 (18 Zilhicce 1326), p. 1; no. 192, February 11, 1909 (20 Muharrem 1327), p. 1; no. 193, February 12, 1909 (21 Muharrem 1327), p. 1.

<sup>20</sup> *Tanin*, no. 160, p. 1. For statements about the pressure on the sultan see *Fikir Hareketleri*, no. 100, p. 343 (see n. 13 above). For affirmation of visits by the CUP to the sultan, see also Abbott (1909, p. 135).

<sup>21</sup> The anchoring of a warship in front of Yıldız Palace is a topic that appears recurrently in reports of early stages of the revolution. For one of a variety of descriptions see Buxton (1909, p. 128).

only in larger cities, but in the lower administrative levels (*nahiye*) as well, assuming many state functions, such as levying taxes and administering the affairs of civilians. Of course, at the time the CUP denied all such charges, but foreign eyewitnesses and the Liberals, the CUP's most ardent critics and one-time allies, presented a rather detailed picture of the CUP presence within the state and their quasi-governmental organization and activities. In a famous and oft-repeated criticism of the Liberals, the CUP had become a "government inside the government."<sup>22</sup> In a few rare instances during the early days, even the CUP's main revolutionary organ alluded to the committee's strong presence within the administration, but only to complain that their control of the state institutions was not complete.<sup>23</sup> The later Young Turk accounts were more forthcoming in confirming the Liberals' claims, as they openly confess to the CUP's extensive presence within the state and to their party's organizing activities in the remotest locations from the earliest days.<sup>24</sup>

Having a firm footing within the administration and the army, the CUP did not feel the need to establish a militia. The Chamber of Deputies, dominated overwhelmingly by CUP members,<sup>25</sup> could rely on the CUP's clout to intimidate the government and to pressure it to carry out the proposed reforms. As we will see, however, the CUP's control over the state organs was still far from complete.

### The National Assembly in Iran

In contrast to the Ottoman constitutionalists, who initially had a strong presence within the state and who used the legal period following stage 2 to further consolidate their position, the Iranians initially lacked any substantial support from within the state and could not make much headway into the state during the period of legal activity. Instead, Iranian society, in comparison to the Ottoman Empire, was characterized by a stronger civil sphere and a far less powerful state. Here, the revolutionaries' most glaring area of success was with the committees, the quasi-governmental institutions they organized in competition with the state.

<sup>22</sup> *İkdam*, no. 5314, March 12, 1909 (19 Safer 1327), p. 1; no. 5324, March 22, 1909 (29 Safer 1327), p. 1; no. 5326, March 24, 1909 (2 Rebiyulevvel 1327), p. 1. See also C. R. Buxton (1909, p. 20) and N. Buxton (1909, pp. 23–24).

<sup>23</sup> *Tanin*, no. 108, November 19, 1908 (24 Şevval 1326), p. 1.

<sup>24</sup> *Tanin*, March 3, 1912 (14 Rebiyulevvel 1330), p. 1; *Fikir Hareketleri*, no. 77, April 11, 1935, pp. 389–91, esp. p. 390; no. 99, September 14, 1935, pp. 325–27, esp. p. 326.

<sup>25</sup> On the CUP's influence over the chamber and its success at blocking the effort to form opposition parties inside the parliament see *BDFA* (1909, p. 105; 1910, pp. 109–14).

In the absence of backing inside the state, these institutions became the most important extraparlimentary source of support for the Iranian National Assembly.

With the beginning of legal activity in stage 2, the Iranian revolutionaries could now form open political associations and publish previously prohibited political newspapers and journals. If Cairo, Istanbul, London, and Calcutta were hotbeds of reformist journalistic activity in the preconstitutional era, now Tehran, the seat of the central government, was the center of new radical newspapers. This period also witnessed the formation of two kinds of revolutionary organizations: official committees and popular committees.<sup>26</sup>

The official committees, in theory if not in practice, were under the command of the National Assembly. To compensate for its lack of support inside the state, the assembly approved ambitious legislation that conferred many critical rights and duties to the official committees on May/June 1907 (Rabi' II 1325). This law allowed the assembly to use the official committees for bolstering its power over the local rulers, the provincial administrative offices, and the governors. For example, the official committees were bestowed with the legal authority to supervise the government's administrative offices, monitor law enforcement, issue warnings, propose reforms for the security and progress of the provinces, investigate complaints concerning the governors and subprovincial heads, inspect the collection of taxes and settle complaints about taxes, investigate requests for tax reduction, and discern damages caused by natural calamities. Furthermore, in certain localities, the official committees could levy new taxes to raise revenue for public works.<sup>27</sup>

The newspaper of the official committee in Tabriz, the provincial capital of Azarbaijan and a hotbed of committee activities, indicated that the committees had established an organizational network that was national

<sup>26</sup> References to the operations of official committees in *Anjuman* and the minutes of parliament clearly demonstrate that many official committees were well established prior to the laws that made their operations legal. For the laws with regard to provincial (*iyalati*), subprovincial (*vilayati*), and municipal committees see *Musavvat I-II*, pp. 64–84, 85–97; see also Browne (1910, pp. 244–45). The distinction between the official and nonofficial committees, however, was not always clear; in many places this distinction broke down almost completely. In Tabriz, e.g., the official and popular committees were composed predominantly of the guilds and trades, and the various popular committees reported to the official committee as their central body (Lambton 1963, p. 46). Thus, many newspapers adopted and referred to them as one unit. See, e.g., *Habl al-Matin*, no. 222, February 4, 1908 (1 Muharram 1326), p. 1.

<sup>27</sup> These legislations are collected in *Musavvat I-II*, pp. 73–77 (see n. 26 above). The law with regard to provincial and subprovincial committees has 122 articles. These are just a few examples selected from articles 67, 68, 87, 89, 91, and 92 to demonstrate the newly acquired power of official committees, and by extension of the assembly, against the provincial governments.

in scope. As required by law, the Tabriz committee had regular exchanges with committees in other cities and with lower administrative districts of Azarbaijan. But, in addition, it kept abreast of developments in other provinces and maintained regular contact with official committees in provincial centers and in large cities; the committee assisted them in time of need and received help when needed.<sup>28</sup>

The official committees stepped beyond the extensive responsibilities legally assigned to them and took many more administrative functions of the provinces into their own hands. They not only kept track of local government administrators and carried out a variety of public works projects, but they also actively responded to the public's grievances. With their strong sense of "popular justice," they interfered in many local governmental functions to restore to the public "rights" violated by the old regime. The inhabitants of the city and of the province at large brought their grievances to the local official committee meetings, which in some locations—like Tabriz—convened in a public space. Typical among these were complaints about government officials, local magnates, or commodity prices, especially that of bread. The members, after deliberation, but always in sympathy with the aggrieved party, reached decisions on how to deal with particular grievances.<sup>29</sup>

In the capital, where by definition official committees could not exist, the popular committees set up among themselves central bodies that gave them cohesiveness and organization that helped make their actions consistent across groups (Kirmani 1972, pp. 47–48). Estimates of the actual number of Tehran popular committees during the first assembly range from 100 to 180.<sup>30</sup> The largest of these, the radical committee of Azarbaijan, had over 2,900 members (Lambton 1963, p. 47). Various sources suggest that despite variation in the makeup of members in different cities and in different committees, a great majority of the popular

<sup>28</sup> The Tabriz committee was in touch with the committees in Tehran, Rasht, Anzali, Isfahan, Shiraz, Qum, Mashhad, Qazvin, and Kirman, among other major cities. As the provincial center of Azarbaijan, it was in regular contact with other subprovincial committees such as those in Maku, Khuy, Salmas, and Urumiyah.

<sup>29</sup> For these and the committee's severe action against the magnates and government officials see *Anjuman*, no. 42, February 13, 1907 (29 Dhu al-Hijja 1324), p. 4; no. 46, February 19, 1907 (6 Muharram 1325), p. 1; no. 53, p. 3; no. 69, pp. 11–14; no. 74, p. 3. For prices see *Anjuman*, no. 38, February 2, 1907 (18 Dhu al-Hijja 1324), pp. 1–2; no. 39, February 4, 1907 (20 Dhu al-Hijja 1324), pp. 3–4; no. 41, February 9, 1907 (25 Dhu al-Hijja 1324), p. 2; no. 42, February 13, 1907 (29 Dhu al-Hijja 1324), p. 4; no. 47, February 26, 1907 (13 Muharram 1325), p. 2; no. 62, April 4, 1907, (20 Safar 1325), p. 4.

<sup>30</sup> The largest estimate is that of Kasravi (1951, p. 569) and the lowest, Lambton (1963, p. 47). Hidayat (1982, pp. 151, 159) and Taqizadah (1957, p. 44) put their numbers at 130 and 140.

and perhaps also the official committees was drawn from various guilds.<sup>31</sup>

Even though never permitted by law, the committees throughout this period were openly armed. From an early date, in the winter of 1906–7, the provincial committee of Tabriz began organizing a popular militia. The Tabriz committee used the increased disorder in the city and the province as a legitimate excuse to patrol the city neighborhoods, muster more resources, build stronger organizations, and gradually to dominate the city (Kasravi 1951, pp. 394–99). The highly visible rituals and military drills of provincial committees had greatly worried the old regime, and they complained about it to the assembly.<sup>32</sup> Equally worrisome for the government were the popular committees in Tehran that followed the lead of committees in Tabriz, Rasht, and Anzali. The difference was that the Tehran committees, no doubt under the influence of the young revolutionary intelligentsia, self-consciously modeled their militia after the French National Guard, and, much to the distaste of moderate constitutionalists who despised the idea, used the French appellation or various Persian translations of it.<sup>33</sup> Initially, the committees advocated the formation of the National Guard under the pretext of national defense.<sup>34</sup> With increased counterrevolutionary activities,<sup>35</sup> however, they openly professed that the real purpose of the National Guard was to defend the constitutional system against internal enemies.<sup>36</sup>

<sup>31</sup> This is a conclusion that is also shared by Lambton (1963, p. 50). An indication was that the popular committees of Tehran, in addition to names that indicated the regional affiliation of their members, bore names such as the Committee of Shoemakers, the Committee of Hatters, of carriage drivers, of booksellers, etc. (Kirmani 1972, pp. 47–48). For another indication of the close relation between guilds and popular committees see *Majlis* 1, August 17, 1907 (8 Rajab 1325), p. 250; August 21, 1907 (12 Rajab 1325), pp. 256–57. For some of the many examples of guilds' presence in official committees see *Anjuman*, no. 72, May 1, 1907 (18 Rabi' I 1325), p. 2; no. 75, May 6, 1907 (23 Rabi' I 1325), pp. 1–2.

<sup>32</sup> *Majlis* 1, May 10, 1907 (27 Rabi' I 1325), p. 163, *Anjuman*, no. 67, 19 April 1907 (6 Rabi' I 1325), p. 3; no. 75, May 6, 1907 (23 Rabi' I 1325), pp. 1–2; Kasravi (1951, pp. 234–37).

<sup>33</sup> They variously called themselves Garde Nationale, Gard-i Milli, Quva-yi Milli, or Nizam-i Milli. For one of many examples see *Habl al-Matin*, no. 101, August 26, 1907 (17 Rajab 1325), pp. 4–5. For the opposition of some constitutionalists see Kirmani (1972, pp. 81–82). See also Dawlatabadi (1983, 2:202–3).

<sup>34</sup> *Habl al-Matin*, no. 157, November 9, 1907 (3 Shawwal 1325), p. 1.

<sup>35</sup> In the face of disorders that signaled the coming of counterrevolution, the committees pressured the assembly to create an official National Guard. Some assembly members openly threatened the inciters of disorder by appealing to the National Guard, in spite of its quasi-legal nature. See *Majlis* 1, November 16, 1907 (10 Shawwal 1325), pp. 385–86; November 18, 1907 (12 Shawwal 1325), pp. 389–90; November 20, 1907 (14 Shawwal 1325), pp. 390–91.

<sup>36</sup> *Habl al-Matin*, no. 222, February 4, 1908 (1 Muharram 1326), pp. 1, 3; no. 223, February 5, 1908 (2 Muharram 1326), pp. 1–2; no. 236, February 25, 1908 (22 Muhar-

That the assembly never legally approved the National Guard did not prevent the committees from forming it. Dressed in special uniforms and using the expertise of a few army officers, they began to drill and organize. And in part to cover the expenses for obtaining arms and setting up the neighborhood militia, they extorted money from the wealthy (Kirmani 1972, pp. 84–86). Estimates of the number of armed committee men, even those made by the contemporaneous observers, vary fantastically, ranging from 2,000 to 100,000.<sup>37</sup> Yet, despite the discrepancy in the reported numbers, all eyewitnesses agreed that the shah, his court, and the government all feared the committees, an institution that was a serious threat to the small and badly equipped Iranian army.

### The Russian Duma's Struggle

The Russian constitutionalists proved to be the weakest of the three groups. Unlike their Ottoman counterparts, they did not command the loyalty of a large section of the armed forces or the bureaucrats. On the other hand, the Russian opposition's impressive ability to form organizations of various kinds during the first few weeks after the October Manifesto proved that Russia's capability for organized opposition surpassed that of Iran. Yet, the Russian constitutionalists' misfortune was that, unlike the Iranians, they confronted a powerful state with the backing of a modernized army. The autocracy, alarmed at the disturbances and the growth of organized opposition following the partial removal of legal restrictions, soon clamped down on these organizations. They also moved to eliminate newly won legal freedoms before the elections could be held for the Duma. These events marked a critical moment for Russian constitutionalism and was its main source of weakness. Once the Duma was established, it was left without extraparlimentary means to coerce the autocracy into accepting its programs for political and social reform.

In Russia the October Manifesto began a crucial period of six weeks (the Days of Freedom) that lasted until early December. An immediate result of the newly won civil freedoms was a proliferation of revolutionary newspapers, magazines, pamphlets, brochures—in short, all types of publications. Although the government officially retained, at least until late November, the right to censor prepublication material, the St. Petersburg soviet on the day following the issuance of the October Manifesto, declared the end of censorship and ordered the workers to refuse

---

ram 1326), pp. 2–3. These pages are replete with references to the French Revolution. See also Dawlatabadi (1983, 2:201–3).

<sup>37</sup> The larger estimate is that of Kirmani (1972, p. 84). Dawlatabadi (1983, 2:203) puts the figure at 2,000.

to print material that had passed through the censors' hands. Thus, the Russian press underwent a spectacular transformation when what had only appeared underground could now be published in abundance and without restriction, making possible the ruthless assailing of government officials by a hostile, public press (Ascher 1988, pp. 231, 276; Harcave 1964, pp. 212–15; Healy 1976, pp. 20, 57–58).

The newly won freedoms also allowed the population to convene, to organize meetings, and to establish legal political parties. As a result, both the number of general meetings, the range of activities, and the membership of political parties of various creeds increased as they found a far more hospitable environment. In the first four weeks following the October Manifesto, more than 400 meetings took place. The liberals were no longer forced to act within the government-imposed restraints and even the socialist parties were provided with a better opportunity to operate; their leaders could now return from Europe with a reduced risk of arrest (Ascher 1988, pp. 276–77; Harcave 1964, pp. 224–25; Healy 1976, p. 59). The Constitutional Democrats (Kadets), the liberal party, was the most important party of this period and occupied a centrist position that leaned to the left. To the left of the Kadets were various socialist parties, the most important of which were the Social Democrats and the Socialist Revolutionaries; slightly to the right of the Kadets was the Union of October 17 or the Octobrists. The largest and most influential of the right-wing parties was the Union of Russian People (Ascher 1988, pp. 234–42; 1992, pp. 31–47; Levin [1940] 1966, pp. 29–34).

The most impressive gains with respect to the revolutionaries' resources were made by labor organizations. It was during the Days of Freedom that unionization gained momentum and labor was provided with an opportunity unmatched in its history. This period saw the beginning of widespread expansion of unions in cities and towns previously untouched by unions and among occupations formerly not unionized. In Moscow alone 67 were established; in St. Petersburg, 58, the great majority of which were founded during November 1905 alone. These unions were united under the leadership of the Central Bureau of Trade Unions (Bonnell 1983, pp. 122–27; Ascher 1988, pp. 242, 276–77; Harcave 1964, p. 215). The most significant strides were made by the soviets, with nearly 50 of them operating under the leadership of the influential and imposing St. Petersburg soviet. The Moscow soviet, which was next in importance to the St. Petersburg soviet, had 80,000 workers as members. In addition to the ties that soviets established among themselves, they extended their network to the labor unions and factory committees to create a form of national organization of labor. Their gains in power were impressive and they were even able to create their own militia. By mid-November, the St. Petersburg soviet claimed 6,000 militia armed

with some type of weapon for the avowed purpose of protecting their meeting places (Bonnell 1983, pp. 125–26, 171–80; Ascher 1988, pp. 219–22, 276–78; Harcave 1964, pp. 212–15, 224–26). Reflecting the conservatives' fear of the growing power of the soviets, Ascher notes that a like-minded newspaper complained that "there were really two governments, one led by Count Witte [the prime minister] and one by Khrustalev Nosar (chairman of the Petersburg soviet), and that no one knew who would arrest whom first" (1988, p. 278).

The waning authority of the autocracy also gave way to defiance by the peasantry. The Days of Freedom saw the first large-scale agrarian revolts since the 18th century with an added ingredient that set it apart from previous uprisings of peasants. For the first time in the history of the empire, peasant uprisings had a political organization that acted in concert with disturbances in the urban areas (Harcave 1964, pp. 216–20; Verner 1990, pp. 105–6; Ascher 1988, pp. 267–69).

The labor radicalism and militancy reached a peak during the Days of Freedom. Blinded by their recent success in winning the October Manifesto, the leaders of the labor movements viewed their victory as a sign of government weakness and engaged in a variety of provocative activities, ranging from a call for a general strike in support of an eight-hour working day, to the soviets' interference with governmental authority in large cities and even its replacement of local governments in some of the outlying regions of the empire. In November, the soviets and various socialist parties were calling for "armed struggle against the tsarist regime." These actions greatly alarmed the government, united the employers against labor, and placed the liberals who did not want to defend the government in a precarious position (Ascher 1988, pp. 275–98).

On November 26 the government took decisive action by arresting the leader and several deputies of the St. Petersburg soviet. In retaliation, on December 2, the newly elected leadership of the St. Petersburg soviet published the provocative Financial Manifesto, a declaration that asked the people to stop making payments to the treasury, to demand payments in gold, to withdraw all deposits from banks in gold and other similar measures, to "cut the government from the last source of its existence: financial revenue." When members of the soviet's executive committee were arrested in retaliation for publication of the Financial Manifesto, the Moscow soviet assumed leadership. But, deprived of the leadership of the central soviet, the Moscow soviet could not lead a successful wave of strikes such as those that had led to the granting of the October Manifesto. In the bloodiest confrontation of the entire movement, the small numbers of inadequately trained and scantily equipped militia of labor proved to be a poor match for the modern Russian army (Ascher 1988,



pp. 279, 298–301, 304–23; Harcave 1964, pp. 232–39; Bonnell 1983, pp. 195–97). Following the suppression of the workers, the government set out on a punitive campaign in the countryside to suppress the peasant uprisings that were spreading with rapid speed (Manning 1982, pp. 141–76; Ascher 1988, 267–68, 330–35; Harcave 1964, pp. 228–30, 240–42). Within four weeks, the government's authority had been restored to levels it had enjoyed before Bloody Sunday. Thus came the first wave of counterrevolutionary backlash that ended the Days of Freedom before the establishment of the Duma, a critical event that severely diminished the extraparlimentary support the Duma could have enjoyed from the organizations of labor.

Before the Russian Duma was established, the soviets were suppressed, their militia crushed, various organizations of labor banned, and civic freedoms curtailed. While the CUP-dominated Ottoman Chamber of Deputies relied on the CUP's supporters inside the state and its semi-clandestine party branches and clubs to intimidate the state, and the Iranian National Assembly drew power from the committees—all in an atmosphere of relaxed legal restrictions—the Russian Duma was deprived of crucial extraparlimentary support to confront its government while the rigid legal restrictions previous to the issuing of the October Manifesto were reimposed. The Russian Duma was thus, from inception, a weaker institution than its counterparts in the Ottoman Empire and Iran.

### CONSTITUTIONAL BATTLES, POLITICAL AND SOCIAL REFORMS, AND COUNTERREVOLUTION

A notable change during stage 2 was that now a large part of the battle over state power was fought within the legal framework of the Russian, Ottoman, and Iranian representative assemblies. These radical assemblies were not merely another legal resource in the hands of the challengers; they qualitatively transformed the dynamics of the fight for the capture of state power.

Constitutional revolutionaries entered battles with the old regimes on two fronts. First, they strove to transform the political structure of the old regimes and to make the assemblies the dominant force in government. In a language that bore striking similarity in all three settings, the elected representatives debated central issues such as who or which institution had legislative powers, the right to override assemblies' decisions, the authority to introduce new bills, the right to appoint the head of the cabinet and individual cabinet ministers, the extent of the cabinet's and each minister's responsibility toward the assembly, and the locus of sovereignty (whether it belonged to the nation or the monarch). This period

saw a profusion of interpellation, the dismissals of individual ministers, and the fall of entire cabinets. The second aspect of this fight was over the assemblies' intention to radically intervene with sweeping social programs. In the Ottoman Empire and in Iran, the constitutionalists' program included a complete revamping of the states' administrative structures as well.

What determined whether the legal parliaments had the ability to enforce their decisions was the support they received from extraparliamentary and illegal sources. In the Ottoman Empire, these included the support of military officers and administrative bureaucrats as well as the CUP party structure and clubs; in Iran these sources consisted of the official and popular committees. In Russia, the soviets and the zemstvos (elected local committees introduced in 1864) could have played an equivalent role by lending support to the Duma and pressuring the government to abide its orders. Yet, the soviets were suppressed at an early phase and the zemstvos, which had initially supported the constitutionalists' cause, drastically changed their stance in favor of the government. Such support was essential for the assemblies to win constitutional battles with their governments, to modify the constitution to their own advantage or impose their own interpretation of it, to force the appointment of sympathetic ministers and prime ministers, and to not only suggest radical reforms but to begin implementing them. On all of these counts, the Ottoman Chamber of Deputies, which had greatest access to extraparliamentary sources of support, was the most successful in winning constitutional battles and implementing reforms. The Iranians came next, and the Russians, who at an early date were deprived of all extraparliamentary sources of support, came to have the most feeble of all assemblies and could not implement any of their central reform programs.

While the exact content of these programs varied in each setting, their effect was highly destabilizing in all. And if compromises with the old regimes were forged and revolutionary coalitions were formed over ambiguous goals, the assemblies left no room for ambiguity in their intentions. On the one hand the monarchs and governments realized the constitutionalists strove to make them subservient to the assemblies, while a wide array of social classes understood that if the demanded reforms were implemented they would fall victim to the very assemblies they helped establish. These developments led to more intensified fights with the old regimes and to the breakdown of prerevolutionary coalitions, prompting many to actively side with the old regimes they had previously opposed.

There is nothing surprising about introducing radical reforms after the revolution. Nor is there anything surprising about the negative reaction

reforms stir among the previously apathetic or sympathetic social classes. What should be highlighted are the different settings within which the socialists and constitutionalists introduced their programs and the comparative disadvantage of the constitutionalists in this context. In contrast to the socialists, who introduced radical programs after capturing state power in a situation where they were, formally at least, in absolute command of the state, the constitutionalists introduced their programs while still struggling with the old regimes over state power. Consequently, constitutionalists, unlike socialists, could not rely on inherited state power to check the negative reaction to their reforms. The constitutionalists began their reforms with, at best, an unstable power base, which allowed the negatively affected social groups to find ready allies with the old regimes still in command. Not being in full control of the state administration and its coercive organs when they began to introduce their programs, the constitutionalists faced a counterrevolutionary backlash that forced all of them out of power.

### The Sultan's Sovereignty Challenged

The conflict over the constitution in the aftermath of the Young Turk Revolution occurred almost immediately and well before the Ottoman Chamber of Deputies had convened, for in the Ottoman Empire a constitution had existed since 1876.<sup>38</sup> The sultan used the opportunity brought about by the commotion of the early days of stage 2 to issue an imperial decree (August 2, 1908) and constitutionally transfer the right to appoint the ministers of the army and the navy to himself. The constitution of 1876 had left open whether these ministers were to be chosen by the sultan or the grand vizier, and the newly issued decree made this the sultan's prerogative. In response, the CUP challenged the sultan, arguing that changing the constitution was a right of the chamber, and since it had not yet convened, it was up to the grand vizier to appoint the ministers.<sup>39</sup> On the other hand, the CUP used this occasion to force the complaisant Grand Vizier Said Pasha out of power after criticizing him and his cabinet for their old-fashioned style and their incompatibility with

<sup>38</sup> For the granting of the constitution and the Young Ottoman movement, see Mardin (1962) and Davison (1963).

<sup>39</sup> For the sultan's decree see *Düstur* 2, I, no. 8, August 2, 1908 (4 Recep 1326), pp 11–14. For CUP's criticism see *Tanin*, no. 2, August 3, 1908 (5 Recep 1326), p. 3; no. 3, August 4, 1908 (6 Recep 1326), p. 1; no. 4, August 5, 1908 (7 Recep 1326), pp 1–2, esp. p. 1. See also Ahmad (1969, p. 19).

a modern constitutional government, which required responsibility and accountability.<sup>40</sup> With pressure on the government, the CUP was not only able to dismiss the Grand Vizier Said Pasha within two weeks after the establishment of the constitutional system, but it was also promised that ministers of the army and the navy, like all other ministers, would be appointed by the new grand vizier rather than the sultan.<sup>41</sup>

When the CUP became disgruntled with the subsequently appointed Grand Vizier Kamil Pasha, he and individual ministers were also criticized for their old-fashioned behavior and ignorance of the constitutional system. In its new round of attacks against the government during Kamil Pasha's reign, the CUP began to challenge even the explicit clauses of the constitution of 1876 that bestowed upon the sultan the right to choose the grand vizier, arguing that this right belonged to the chamber and, by extension, the "nation," or the locus of sovereignty.<sup>42</sup> Tensions reached a new height when the grand vizier changed three cabinet ministers, including those of the army and navy, without receiving a vote of confirmation from the chamber. This act was interpreted as an encroachment upon the rights of the chamber and brought accusations that likened Kamil Pasha to the whimsical grand viziers of the old regime.<sup>43</sup> As discussed above, the CUP, through its internal machinations, had already secured guarantees for Kamil Pasha's dismissal, even though he was finally dismissed by the chamber through constitutional procedures. For the third grand vizier, the CUP was able to impose the choice of Hilmi Pasha—a candidate it found to be in agreement with its views—on February 14, 1909. Yet, despite accusations by their critics to the contrary, the CUP was still far from controlling the cabinet and continued to have disagreements with the last grand vizier as well.<sup>44</sup>

The battles between the Ottoman government and the Chamber of Deputies were not confined to their respective constitutional rights and

<sup>40</sup> *Tanin*, no. 5, p. 3 and pp. 3–4 (see n. 18 above).

<sup>41</sup> *Fikir Hareketleri*, no. 96, August 24, 1935, p. 277. For announcement of Said Pasha's resignation see *Tanin*, no. 6, August 6, 1908 (9 Recep 1326), p. 1; and Ahmad (1969, p. 20).

<sup>42</sup> *Tanin*, no. 55, September 25, 1908 (28 Şaban 1326), p. 1; no. 57, September 27, 1908 (30 Şaban 1326), p. 1; no. 128, pp. 1–2, no. 129, pp. 1–2; no. 145, p. 1; no. 157, p. 1, no. 160, p. 1 (see n. 19 above).

<sup>43</sup> *Tanin*, no. 192, p. 1; no. 193, p. 1 (see n. 19 above).

<sup>44</sup> *Tanin*, no. 194, February 13, 1909 (22 Muharrem 1327), p. 1; no. 216, March 8, 1909 (15 Safer 1327). For the Liberals' accusation that the CUP had created a government according to its will, see *İkdam*, no. 5322, March 20, 1909 (27 Safer 1327), p. 1. For rather minor disagreements between the CUP and the new government, see *Tanin*, no. 235, March 27, 1909 (4 Rebiyülevvel 1327), p. 1.

political power. They also involved the sensitive issue of administrative reform. In fact, the counterrevolution that occurred 10 months into the second stage was a direct reaction to the rationalization of staff policies and purges.<sup>45</sup> The groups that coalesced and staged the short-lived counterrevolutionary movement, especially the purged, demoted, displaced, or threatened members of the civil and military bureaucracy, attempted to reverse the tide of the CUP's staff reforms.

During the four months that the Ottoman Chamber of Deputies was in session (December 17, 1908–April 12, 1909), it approved an impressive amount of legislation. The most important bills concerned the staff reorganizations within the civil bureaucracy and the army. The CUP commanded many followers among civil bureaucrats and military personnel, and, after the revolution, the military and civil bureaucracy became the site of extensive rationalizing transformations aimed at ending the bureaucratic dichotomy between the modern and patrimonial soldiers and officers. In the eyes of the CUP, a major problem the army faced was the inflated number of high-ranking officers and the extremely young age of many high-ranking military commanders who had gained their promotion through connection to a patron or the sultan. To correct this situation, the CUP recommended purges, demotions, and early retirement for many officers to replace them with educated ones.<sup>46</sup> The same solution was suggested for the civil bureaucracy, which was plagued with similar conflicting rules of operation. The CUP-dominated Chamber of Deputies, by approving legislation to purge officials within civil bureaucracy and the military, gave the upper hand to officials who had modern training and were advocates of rational administrative procedures.<sup>47</sup>

<sup>45</sup> In discussions so far, this aspect of the counterrevolutionary has always been underemphasized at the expense of the superficially religious form that it took. For a contrasting view of this incident that describes it as essentially motivated by soldiers' interests in a religious guise see Akşin (1971, pp. 309, 336).

<sup>46</sup> *Tanin*, no. 18, August 19, 1908 (21 Recep 1326), pp. 1–3. A British official was struck by the large number of officers in the Ottoman navy, an organization that was, in his opinion, "virtually nonexistent." In his report he indicated that there were 7,500 officers in the navy's active list, compared to Britain's 5,000 (*BDFA* 1908, p. 22).

<sup>47</sup> The major legislation concerning purges came during the fourth revolutionary stage and after the CUP dominated the state entirely. For the central legislation on purges within civil officialdom during stage 2, see *Düstur2*, I, August 15, 1908 (17 Recep 1326), pp. 55–56. Other purge legislation prior to the counterrevolution was as follows: *Düstur2*, I, no. 10, August 12, 1908 (14 Recep 1326), pp. 39–40; no. 20, August 22, 1908 (24 Recep 1326), pp. 61–62; no. 21, August 22, 1908 (24 Recep 1326), pp. 62–63; no. 26, September 3, 1908 (6 Şaban 1326), pp. 71–72; no. 34, October 3, 1908 (7 Ramazan 1326), pp. 85–88. For the army, see *Düstur2*, I, no. 87, June 27, 1909 (8 Cemaziyelahir 1327), pp. 324–25.

The extensive staff reorganization and purges were clearly disruptive for the displaced members of the traditional bureaucracy and the army. Their negative reactions to the reforms culminated in a sudden and unexpected counterrevolutionary uprising that began within the army, where the reforms had hit the hardest. In the early morning of April 13, 1909, soldiers of the light infantry battalions of the Third Army Corps (*avcı taburlar*), after congregating in large numbers at the square of Sultan Ahmed and Ayasofya without their officers, chose a sergeant as their leader—an action with clear symbolic significance—and presented the *şeyhülislam* (the highest religious authority) with five demands.<sup>48</sup>

Whereas the soldiers' third demand was the restoration of shari'a (*şeriat*, religious law) requested under the influence of the religious groups with whom they cooperated,<sup>49</sup> the most significant of the soldiers' demands was the fourth, which called for "banishing and replacing their superior educated officers [*mektepli*] and reappointing the officers who had risen through ranks [*alaylı*, or old troopers] and were wrongfully harmed by being fired."<sup>50</sup> The fourth demand clearly revealed that the old troopers (*alaylı*) were reacting to the CUP's policy of privileging the officers with modern education (*mektepli*), that is, the CUP's supporters.

The old trooper officers of the First Army Corps who joined the rebelling infantry soldiers also protested against the staff policies of the CUP. In its efforts to reduce the influence, significance, and numbers of the old trooper officers, the CUP had fired 1,400 officers from the First Army Corps.<sup>51</sup> The numerically and politically more significant soldiers of the First Army Corps, whose officers had mostly risen through the ranks, were favored by the sultan: they served as his personal military guards and enjoyed excellent material conditions (Farhi 1971, p. 281). The soldiers of the First Army, after arresting their educated superiors and killing several of them, joined the rebelling infantry soldiers of the Third

<sup>48</sup> See Nadi (1909, pp. 33–34, 39), Cevat (1960, pp. 88–89), and *İkdam*, no. 5347, April 14, 1909 (23 Rebiyulevvel 1327), p. 1.

<sup>49</sup> For a competent description of the counterrevolution and its demands—though one that gives overwhelming import to the third demand—see Farhi (1971). For views that give greater importance to the religious aspect of the movement see Shaw and Shaw (1977, pp. 279–80) and *BDFA* (1908, pp. 22–23). For one that underplays religion, but without making central the issue of bureaucratic dichotomy, see Ahmad (1969, pp. 40–45).

<sup>50</sup> See Nadi (1909, p. 36). For a full set of demands see Nadi (1909, pp. 35–36) and Farhi (1971, pp. 275–76).

<sup>51</sup> Akşin (1971, pp. 46–47). Knight (1909, pp. 329–30) perceptively reports on the grievances of the First Army Corps and the division between the *alaylı* and *mektepli* officers. He further distinguishes the demands of the latter from the students of religion and their teachers, who preached in the barracks that the committee was endangering the Muslim faith.

Army and, in a threatening letter addressed to the ministers and carrying the requests of more than 7,500 officers of the various army corps, they demanded the reassignment of purged officers and a halt to such reforms (Nadi 1909, pp. 55–57).

In their gathering, the rebellious soldiers were joined by a variety of groups, most significantly by thousands of teachers of religion (*hocas*), students of religion (*softas*), lower-ranking religious clergy (*imams*), and preachers. Having started their counterrevolutionary activities earlier, the lower-ranking religious groups had officially announced their existence as a political group, the Society of Muhammad (*İttihad-i Muhammedi*). They joined the soldiers to protest the secular policies of the CUP, such as its stance on minorities and the state's encroachment into the domain of religion through various reforms.<sup>52</sup> Paradoxically, the upper-ranking clergy did not join forces with anti-Unionist lower clergy, for the former had developed a strong animosity for Abdülhamid II during his reign and, perhaps more significantly, because they had not yet realized the extent of the CUP's secularist policies and the gravity of the threat it posed to the religious institutions of the empire.<sup>53</sup>

The counterrevolutionary movement was not confined to the old troopers and lower-ranking religious orders. Other groups, such as the Sublime Porte and the Liberals, took active interest in it. The Sublime Porte was represented by the deposed Grand Vizier Kamil Pasha, around whom gathered similar high-ranking officials of the Porte and scores of lower-ranking officials. The latter, like the old troopers, had either fallen victim to the CUP purges or felt imminently threatened (BDFA 1910, p. 110; Danişmend 1961, p. 22; Farhi 1977, p. 280–81; Knight 1909, p. 328). The Liberal opposition was represented by Prince Sabahaddin who, as a staunch opponent of Abdülhamid, had cooperated with the CUP in Europe under the broad-based coalition of the Young Turks but was now opposing the CUP over its centralizing policies. Sabahaddin was the champion of Anglo-Saxon laissez-faire ideology and was steadfastly opposed to the French centralization model the CUP espoused; in his view, this model had led to the replacement of the Hamidian dictatorship

<sup>52</sup> The secular policies of the committee were directed at reducing the power of the clergy, especially the clergy's influence over judicial matters. For the participants and the minority question, see Farhi (1971, pp. 275, 281), Shaw and Shaw (1977, pp. 279–80), and Ahmad (1969). For the political manifesto and organizational framework and the speech by Derviş Vahdeti in which he announced the formation of the group, see Tunaya (1984, pp. 199–205) and Danişmend (1961, p. 22).

<sup>53</sup> During the counterrevolution, the orthodox ulema condemned the sultan for persecution of religion during his rule and declared their support for the CUP (Nadi 1909, pp. 62–64). When they did realize the extent of the threat and withdrew their support it was already too late to stop the further encroachment of the CUP.

with that of the CUP (see Sabahaddin 1908*a*, 1908*b*, esp. pp. 28–29, 41–43; Kuran 1945, 1948). In order to further their own ends against the CUP, the various opposition groups, in spite of clearly conflicting goals, cooperated with one another and especially with the religious group headed by Derviş Vahdeti. Some even claim that, in the political vacuum created by the CUP's retreat to Salonika, the Liberals stepped in to assume political power (see Shaw and Shaw 1977, p. 280; Tunaya 1984, pp. 145–53; Ahmad 1969, p. 43; Halid 1909, pp. 758–60; Knight 1909, pp. 323, 328).

The extent of the palace's participation in the counterrevolution is not altogether clear. High-ranking politicians close to the sultan denied his involvement at any point during the 10 days or so that this event was underway (Danişmend, 1961, pp. 18–21, 25–33; Cevat 1960, p. 58; Akşin, 1971, pp. 364–65). On the other hand, while scholars may be divided over the question of the extent of the sultan's initial support, they agree that eventually he supported it and used this opportunity to restore his lost power (Tunaya 1984, pp. 196–97; Shaw and Shaw, 1977, p. 281). A less contested issue, however, was the palace's active participation through the heir apparent (Tunaya 1984, p. 197).

Thus, in reaction to the short-lived counterrevolutionary movement in Istanbul, the CUP temporarily retreated to Macedonia, where the revolution had originated and where the CUP had its strongest organizational basis. The Chamber of Deputies, symbolically, was not officially disbanded and a newly elected Liberal deputy replaced the ousted CUP president.

### The Balance of Power in Iran

The first Iranian National Assembly functioned for a little more than 20 months (October 7, 1906–June 23, 1908). Yet in this short period, cabinets changed not less than nine times.<sup>54</sup> From the early days, the representatives were engaged in a debate over the assembly's power and attempted to endow it with the right to dismiss cabinet members.<sup>55</sup> With assistance from the radical newspapers, the representatives exerted a great effort to construct a constitutional concept of politics and to break away from the traditional notions of statecraft and kingship in the Iranian scene. Particularly suitable occasions for explicating the meaning and functions of a modern state were the interpellations, during which the representatives severely criticized the ministers for their legal infractions,

<sup>54</sup> For a list of the members of these cabinets see Burujini (1971, pp. 1–28).

<sup>55</sup> For the first episode of this fight, see *Anjuman*, no. 48, February 28, 1907 (15 Muharram 1325), p. 2; Kasravi (1951, pp. 215–17).



pointed out that they were responsible to the assembly, and demanded accountability for their actions.<sup>56</sup> The constitutional battles prompted the representatives to draft a supplement to the Fundamental Laws to expand greatly the powers of the National Assembly against the government, an action that created intense friction. The government's resistance to the supplement cost it the prominent Grand Vizier Atabak, whose assassination on August 31, 1907, was backed by the committees, almost all of which considered him an ardent anticonstitutionalist and opposed to the supplement.<sup>57</sup> The supplement that was finally approved on October 7, 1907, marked a victory for the constitutionalists. In contrast to the earlier-ratified constitution, the ministers were now considered responsible to the assembly rather than the shah, and the assembly was given full powers to dismiss individual ministers or to discharge entire cabinets without having to prove legal infractions on their part and, significantly, without having to gain the shah's approval.<sup>58</sup> Despite gaining the right to dismiss the cabinet, however, the assembly could not simply rely on its constitutional powers when confronting the government. To assure that the legal orders of the assembly were followed, the committees pressured the opposition by harassing the ministers and issuing threats against their opponents. The committees even attempted—but failed—to assassinate the shah. Throughout the period of its operation, the first Iranian National Assembly found merely one grand vizier that was to its liking and one cabinet of which it almost entirely ap-

<sup>56</sup> *Majlis*1, January 6, 1907 (21 Dhu al-Qa'da 1324), pp. 46–47; January 13, 1907 (28 Dhu al-Qa'da 1324), p. 50; January 22, 1907 (7 Dhu al-Hijja 1324), pp. 59–60; January 24, 1907 (9 Dhu al-Hijja 1324), pp. 63–64; January 31, 1907 (16 Dhu al-Hijja 1324), pp. 70–72; March 14, 1907 (29 Muharram 1325), pp. 107–9; March 27, 1908 (23 Safar 1326), p. 492. *Habl al-Matin*, no. 1, April 28, 1907 (15 Rabi' I 1325), pp. 3–4. For criticism of newspaper see, e.g., *Habl al-Matin*, no. 53, June 28, 1907 (17 Jumada I 1325), pp. 1–3; *Anjuman*, no. 44, February 14, 1907 (1 Muharram 1325), pp. 1–4; no. 45, February 17, 1907 (4 Muharram 1325), pp. 1–4.

<sup>57</sup> *Majlis*1, August 30, 1907 (21 Rajab 1325), p. 267; August 31, 1907 (22 Rajab 1325), p. 269; Kasravi (1951, pp. 445–50), Browne (1910, pp. 150–51), and Dawlatabadi (1983, 2:140–43); *Habl al-Matin*, October 5, 1907 (26 Sha'ban 1325), p. 2. For mourning ceremonies by the committees commemorating Atabak's assassin, see *Habl al-Matin*, no. 135, October 7, 1907 (29 Sha'ban 1325), pp. 5–6; see also Browne (1910, pp. 151–54) and Kasravi (1951, pp. 464–65).

<sup>58</sup> Even prior to approval of the supplement, the assembly took votes that overwhelmingly approved of dismissing the entire cabinets. See *Majlis*1, April 24, 1907 (11 Rabi' I 1325), pp. 151–53; April 29, 1907 (16 Rabi' I 1325), pp. 153–54. For articles on the powers of the assembly in the original constitution and for changes in the supplement, see *Musavvat I-II*, pp. 8–9, 27–28 (see n. 26 above).

proved.<sup>59</sup> Even this government was soon removed from power by the shah and its prime minister exiled.<sup>60</sup> Yet, thanks to the agitations of the committees, the assembly repeatedly confronted the government with a cabinet crisis and forced the constant circulation or removal of ministers and prime ministers—though they rarely approved of the replacements. The Iranian National Assembly and the government had reached some form of balance of power where severe actions by one were responded to with similar actions from the other.

Even more so than the Ottoman Chamber of Deputies, the Iranian National Assembly commenced reforms at a time when it did not yet dominate the state; it lacked sufficient coercive power to suppress the disturbances the reforms generated. On the one hand, the assembly's financial reforms fomented strong reaction from a wide array of social actors. On the other, the shah and his government considered the constitutional system a threat to the shah's sovereignty. They regarded the assembly's quest for dominating the state a disturbing encroachment upon their age-old prerogatives. Furthermore, the assembly's legislative undertakings together with its judicial and educational reforms contested the authority of the traditional religious institutions. Thus, the constitutionalists in Iran witnessed a daily swelling of the opposition's ranks and found themselves incapable to stem its growth.

The assembly proposed financial reforms to balance the budget by raising taxes for the deficit-ridden treasury of the old regime. These marginally implemented reforms consisted broadly of rationalized budgetary allocation and taxation. Rationalized budgetary allocations included salary reduction for state employees, extensive cutbacks in court and palace budgets, ministerial allocations according to detailed expenditure records, and abolition of a traditional compensatory method called *tuyul*. The rational taxation plan included the replacement of a traditional category paid by landowners as a portion of their taxes and the reorganization of the provincial tax structure to bring it under greater central control.<sup>61</sup>

<sup>59</sup> Majlis was more than ready to use its right. A vote of no confidence was given nine days after the signing of the supplement. See *Majlis* 1, 16 October 1907 (9 Ramadan 1325), p. 349; Burujini (1971, p. 16). For pressure over the signing of the supplement and the shah's anger over the introduction of a pro-National Assembly cabinet, see Dawlatabadi (1983, 2:151–53) and Hidayat (1982, pp. 159–60). For statements of support for the new cabinet see *Habl al-Matin*, no. 152, November 4, 1907 (28 Ramadan 1325), pp. 2–3.

<sup>60</sup> *Habl al-Matin*, no. 199, January 6, 1908 (1 Dhu al-Hijja 1325); no. 202, January 8, 1908 (4 Dhu al-Hijja 1325), pp. 1–2; Browne (1910, pp. 162–63, 165); Hidayat (1982, p. 160).

<sup>61</sup> The rationalized taxation plan was based on auditing, which was in turn dependent upon precise information obtained through a nationwide cadastre. The latter was never carried out. Yet, pressured to solve the chronic financial problems of the old regime, the assembly haphazardly set out to implement extensive financial reforms.

The abolition of the *tuyul* was one of the most critical decisions of the first National Assembly. The category *tuyul* constituted the income from land assigned in lieu of salary from the state.<sup>62</sup> Its abolition, together with other financial undertakings, brought about a strong reaction in the capital and provinces.<sup>63</sup> The response of *tuyul* holders was indeed strong, particularly in view of the fact that, much to the representatives' chagrin, the majority of the assignments continued to remain in the hands of *tuyul* recipients.<sup>64</sup>

The financial reforms also included another sensitive area: salary reforms. The opposition soon exploited this issue. In a representative's words, "malicious individuals" were spreading the rumor about the assembly's intention to reduce the salaries of government officials and dependents, rich or poor. The assembly and the revolutionary newspapers vehemently denied this rumor, arguing that salary reforms were not in-

---

For a discussion of the provincial tax structure and *tafavut-i 'amal*, see *Majlis1*, January 6, 1907 (21 Dhu al-Qa'da 1324), p. 45. Later discussions clearly demonstrated that a consensus on precise working of provincial tax structure did not exist. For *tuyul* and traditional cash payments in place of taxes in kind (*tas'ir*) see Adamiyat (1976, pp. 448–50), Kasravi (1951, pp. 228–29), and Browne (1910, pp. 238–39). The resulting problems prompted the assembly to approve a bill that required all money collected in the name of *tafavut-i 'amal* to come first to the center and from there to be allocated to the provincial centers. See *Majlis1*, April 6, 1907 (22 Safar 1325), pp. 125–26; Adamiyat (1976, pp. 447–48).

<sup>62</sup> As Arjomand (1988, p. 22) and Lambton (1953) have noted, *tuyul* could constitute income from a variety of sources, but the constitutionalists' central concern was with the land assignments. The centralizing constitutionalists intended to reclaim the *tuyul* and to assign to the deserving owners a regular salary from the state. From the assembly's view, not only were such assignments ill-suited to a modern state, but more important, the actual value of many such assignments had much surpassed their nominal value at the time of their grant. For a description of *tuyul* from the assembly's view see *Majlis1*, January 6, 1907 (21 Dhu al-Qa'da 1324), p. 45; March 19, 1907 (4 Safar 1325), pp. 111–12. Furthermore, the land or taxes assigned as *tuyul* tended to become the private property of the assignees, especially in the second half of the 19th century.

<sup>63</sup> In Tehran, the first large-scale meeting against the abolition of the *tuyul* was held by a prominent holder who organized a gathering attended by similarly disgruntled holders. See *Majlis1*, June 6, 1907 (24 Rabi' II 1325), pp. 183–84; *Habi al-Matin*, no. 38, June 11, 1907 (29 Rabi' II 1325), p. 2; no. 49, June 23, 1907 (12 Jumada I 1325), p. 1; *Sur-i Isrâfil*, no. 5, June 26, 1907 (15 Jumada I 1325), p. 3. In addition to Tehran, the assembly's financial decisions—first with regard to *tuyul* and then with regard to *tafavut-i 'amal* and *tas'ir*—created quite a commotion in provincial regions and in every major city, many times in collaboration with provincial governors and rulers. See, e.g., *Majlis1*, August 12, 1907 (3 Rajab 1325), p. 242. The newspapers also reported the disorders caused by landowners and the wealthy in general; see *Musavat*, no. 23, May 11, 1908 (9 Rabi' II 1326), p. 7; *Anjuman*, no. 69, April 24, 1907 (11 Rabi' I 1325), p. 4; no. 70, April 26, 1907 (13 Rabi' I 1325), p. 4.

<sup>64</sup> For indications of the assembly's lack of success see *Anjuman*, nos. 77–78, May 4, 1907 (21 Rabi' I 1325), p. 3; *Majlis1*, August 26, 1907 (17 Rajab 1325), pp. 263–64.

tended to affect the poor, the middle class, or the ordinary state employees but the princes, court members, and individuals with "salaries larger than some foreign presidents."<sup>65</sup> Overall, only 2,000 individuals were directly affected, yet the salary reductions managed to spark widespread disorders. The palace's budget was also reduced, and the assembly chose not to respond to the shah's protest about the insufficiency of the assigned funds.<sup>66</sup>

To organize the counterrevolution, the government supported various groups with money and labor. It was no secret that the shah, his court, and high-ranking government members organized clandestine meetings to unite the opposition, that they were attempting to attract the support of the cities' poor, the holders of *tuyul*, and those threatened by salary cuts, and that they even prepared the army for a military attack on the assembly.<sup>67</sup> In addition, the shah and his government funneled large financial sums to the conservative clergy who, in the third week of June 1907, took sanctuary in protest to the assembly and gathered a large crowd of supporters.<sup>68</sup>

As a lawmaking body, the National Assembly posed a serious threat to traditional religious authority. The late 19th-century legal sphere was already a contested scene as the Qajar state attempted to extend the jurisdiction of the secular customary law (*'urf*) over that of the religious law (*shari'a*; see Algar 1969, pp. 11–13, 128, 169–71, 223–24; Nashat 1982, pp. 43–54). When the assembly began operation in 1906, it entered this contested scene without specifying the limits of its lawmaking activities; its presence carried the threat that it could encroach on the territory of both the traditional secular and religious law and ultimately direct both. The assembly's impact on the religious law, however, was of greater consequence. The assembly intended to legislate new laws, some

<sup>65</sup> *Majlis*1, August 16, 1907 (7 Rajab 1325), pp. 249, 252; *Habl al-Matin*, no. 166, November 19, 1907 (13 Shawwal 1325), pp. 1–3.

<sup>66</sup> *Majlis*1, October 30, 1907 (23 Ramadan 1325), p. 363; November 2, 1907 (26 Ramadan 1325), p. 365; November 9, 1907 (3 Shawwal 1325), p. 376; November 16, 1907 (10 Shawwal 1325), pp. 385–86; November 18, 1907 (12 Shawwal 1325), p. 389; Dawlatabadi (1983, 2:114).

<sup>67</sup> *Majlis*1, April 29, 1907 (16 Rabi' I 1325), pp. 154–55; Kasravi (1951, p. 499); *Musavat*, no. 6, November 23, 1907 (17 Shawwal 1325), no. 9, January 11, 1908 (7 Dhu al-Hijja 1325). In its sixth issue, the newspaper *Musavat* appealed directly to the poor to ignore the instigation of the powerful and the court elements who had found their traditional vested interests threatened by the assembly. For instances when the shah, the government, and the court were directly implicated for organizing the counterrevolution see *Majlis*1, June 6, 1907 (24 Rabi' II 1325), pp. 183–85; June 26, 1907 (15 Jumada I 1325), p. 196; *Sur-i Israfil*, no. 5, June 26, 1907 (15 Jumada I 1325), p. 3; see also Kasravi (1951, pp. 364, 366, 374–75).

<sup>68</sup> Dawlatabadi (1983, 2:129–30); *Sur-i Israfil*, no. 5, p. 3.

directly of European origin. Most significantly, the Supplement to the Fundamental Laws was influenced greatly by the Belgian constitution. As the early unpublished drafts of the supplement clearly indicate, by failing to mention religious courts, the National Assembly intended to end the clerical judicial functions and replace them entirely with secular, state courts (Afshar 1989, doc. 67, pp. 91–100). The assembly and its reforms weighed heavily on the conservative clergy, who considered the shari'a to be a complete code of law interpretable by them alone. The clergy proclaimed itself the sole administrator of justice, and in practice, was bestowed with almost complete judicial responsibility.

Before the final passage of the supplement, the conservative clergy opposed the legislative undertakings of the assembly and the translation of the Belgian constitution after citing the latter's opposition to the laws of religion. It organized an extensive campaign to bring these activities to a halt.<sup>69</sup> It now demanded the creation of a five-member clerical council that was to supervise legislation in order to ensure consistency with religion. The council was to be given the power to revoke legislative bills in cases of deviation (Kasravi 1951, pp. 370–72; Dawlatabadi 1983, 2:108–9).

Despite the assembly's strong resistance to the conservative clergy's proposition, it was finally forced to depart from the radically secular and statist version of the originally intended supplement. As a compromise, the newly added Article 2 conceded to the clergy the right to nominate 20 clerics for the council from whom the deputies would elect at least five members to supervise legislation and to prevent the passage of any law that contradicted religious laws. Thus, even though the assembly agreed to a five-member clerical council, it gave itself some control over the election of its members. Another compromise occurred over the courts of law since the approved version of the supplement, by mentioning the religious courts, sanctioned their continued existence. Yet the supplement left the distinction between the religious and secular courts intentionally vague without substantiating the duties of each in any detail.<sup>70</sup>

<sup>69</sup> *Majlis*1, May 16, 1907 (3 Rabi' II 1325), p. 167; May 21, 1907 (8 Rabi' II 1325), p. 171; June 8, 1907 (26 Rabi' II 1325), p. 183.

<sup>70</sup> For the complete text of the supplement see *Musavvat I-II*, pp. 15–33; see esp. pp. 15–16 (see n. 26 above). See also *Habl al-Matin*, no. 138, October 10, 1907 (3 Ramadan 1325), pp. 2–6. The early, unapproved drafts of the supplement described in much greater detail the duties of secular state courts and unambiguously denied any role for the religious courts. For two critical and early handwritten drafts of the supplement—which did not include many of the later compromises of the assembly, including Article 2—see Afshar (1989, doc. 67, pp. 91–100). For the clergy's initial and later objections to the supplement and the assembly's compromise see also Arjomand (1993).

Even with these compromise offerings, the conservative clerics continued to oppose the supplement on grounds of its threat to clerical institutions. For one, Article 19 made education compulsory and brought all schools (state and private) under the supervision of the Ministry of Education.<sup>71</sup> Traditionally, with the exception of some newly established modern schools, education was the domain of the religious establishment. Furthermore, the clerics were aware of the constitutionalists' intention to abolish the religious courts; the imprecise wording of the final draft of the supplement did not do much to assure the clerics of a secure continued role in judicial affairs. Finally, the conservative clergy rejected the compromise in Article 2 of the supplement on the grounds that no one other than the clergy was eligible to choose the council members.<sup>72</sup>

The opposition of the conservative clergy eventually polarized the clergy as a gradual division developed between the lower and upper ranks of the clerical establishment. Clerical support for the constitutionalists was substantially weakened, for now only a few higher-ranking clerics, some lower-ranking ones who had intimate ties to the guilds, and some religious students remained in support of the constitutionalists.<sup>73</sup> Even the two high-ranking clerics who lent a semblance of religious legitimacy to the assembly were under constant pressure from the newspapers and committees to end the quietist stance they were accused of adopting and to be more impassioned in defending the assembly and the supplement.<sup>74</sup>

Similar to the revolution and counterrevolution in the Ottoman Empire, and in a manner akin to the initial revolutionary movement in Iran, the actors who joined the counterrevolution were not motivated by a single cause. Yet, once again, the counterrevolutionary coalition formulated its opposition to the assembly around a single cause and successfully made the leap from the "particular" to the "general." Almost all the groups in the final coalition had begun their opposition activities independently of the others, but all had eventually adopted the conservative clergy's oppositionist view, the main tenet of which was the defense of Islam against the heathen constitutionalists (Turkaman 1983, pp. 149–51, 240–41, 260–69, 287–88, 296–97, 331). In its final stage, the

<sup>71</sup> *Musavvat I-II*, p. 18 (n. 26 above).

<sup>72</sup> The clergy announced its opposition in a newspaper it published in Shah Abdul'azim (see Turkaman 1983, p. 231).

<sup>73</sup> See Kasravi (1951, pp. 263, 358–60, 375–76, 415–23, 628–29); *Habl al-Matin*, no. 184, December 11, 1907 (5 Dhu al-Qa'da 1325), p. 3.

<sup>74</sup> *Musavat*, no. 4, November 11, 1907 (5 Shawwal 1325), p. 2; no. 9, January 11, 1908 (7 Dhu al-Hijja 1325), pp. 6–7. For the inactivity of Behbahani and Tabatabai in face of counterrevolutionary activities see also Kasravi (1951, p. 263) and Arjomand (1981, pp. 174–90).

oppositionist coalition was composed of members of the court, government employees, palace household employees, major landowners and recipients of *tuyul*, provincial magnates and governors, some small section of the cities' guilds (Dawlatabadi 1983, 2:131), and a large sector of the divided clergy. Like the Ottoman counterrevolutionaries who could not articulate an oppositionist language of their own, the counterrevolutionary coalition in Iran strategically chose the clerics' language to combat the constitutionalists, for it was a ready-made and sophisticated criticism of the constitutionalists that, for the moment, served the opposition well.

The first serious physical confrontation between the constitutionalists and the counterrevolutionaries happened on December 15, 1907, when protestors began a procession that ended at Tupkhanah Square. Here they raised tents and began harassing, robbing, beating, and even hanging bystanders who were suspected of being constitutionalists. Meanwhile, the shah arrested several members of the now proconstitutionalist cabinet and, after the intervention of the British, sent the grand vizier and two other ministers into exile after changing his mind about their execution. This event, which lasted for four days, ended when the shah backed away from a full-fledged attack on the assembly. The crowds dispersed but the assembly was left with a new cabinet handpicked by the shah.<sup>75</sup> When the actual attack came some six months later (June 23, 1908), the armed committees proved to be no match for the modern, Russian-trained Cossack brigade that bombarded the assembly compounds. The assembly fell easily and the constitutional movement seemed to have been defeated (Kasravi 1951, 577–640). In fact, had it not been for the institution-building activities during stage 2, the constitutional movement might have been completely crushed. With the destruction of the National Assembly, the sacking of the constitutional press, and the dispersing of the committees in Tehran, the center of activities moved to Tabriz and the northern Gilan province, where the surviving committees were strongest.

### The Legal Counterrevolution against the Duma

In Russia, the dynamics of counterrevolution differed from those in Iran and the Ottoman Empire in one crucial respect. The Duma was estab-

<sup>75</sup> For a description of events during the Tuupkhanah incident see *Habl al-Matin*, no. 189, December 24, 1907 (19 Dhu al-Qa'da 1325), pp. 1–2, 7; no. 190, December 25, 1907 (20 Dhu al-Qa'da 1325), pp. 1–2; *Musavat*, no. 9, January 12, 1908 (7 Dhu al-Hijja 1325), pp. 4–5; see also Kasravi 1951, pp. 505–12, 521–22; Browne 1910, p. 163; Hidayat 1982, pp. 160–61).

lished after the violent counterrevolutionary backlash that brought a martial end to the Days of Freedom. After the suppression of the soviets, the Duma was left without support from quasi-governmental institutions; thus it was unable to create an effective power bloc for countering the government. In the face of its enfeebled opponent, the government had a free hand to tamper with the terms of its ambiguous compromise and to impose its own definition and render ineffective what the constitutionalists had hoped would be a legislative assembly. Thus began the legal, nonviolent phase of the counterrevolution directed by the bureaucracy and supported by various social classes and institutions. Through a series of legal and electoral interventions, the government significantly reduced the Duma's legislative powers and undermined its representative character by disenfranchising large sectors of the population.

The impotence of the Russian Duma should be compared to the more effective Ottoman Chamber of Deputies. The CUP's extensive presence within the administration and the army provided the Ottoman chamber with the power to dismiss the grand vizier using a minor legal excuse, and the Ottoman revolutionaries proved that they could do so even without employing legal channels. Even the Iranian National Assembly wielded more power than the Russian Duma. Although the Iranian assembly's constitutional authority to dismiss cabinets or ministers was not always respected, the government feared the assembly, for its radical faction enjoyed the strong support of the armed committees. Thus, the government occasionally abided by its orders, and in a few instances it dismissed ministers or even entire cabinets. It was the balance of power between the government and the assembly that permitted the latter to continue its existence in face of many serious legal violations, even after a prominent prime minister was assassinated by members of the committees. In Russia, in striking similarity to the altercations in the Ottoman Empire and Iran, the Duma demanded the subservience of the executive to the legislature; once the executive refused to abide, the Duma asked for the resignation of the cabinet. The similarities, however, end there. As Healy (1976, p. 202) has remarked, the Russian "assembly had no means of compelling an official to answer a query, much less to resign; of course it could not put aside his orders."

The legal machinations of the government for reducing the Duma's powers began with the electoral laws. In July 1905, the government had offered the Bulygin Duma to the opposition, the terms of which reduced the Duma to an essentially consultative body. This offer placed severe limitations on the franchise by proposing a complicated multistage voting system that, combined with a high property qualification (especially in the cities), excluded many, and the industrial workers were disenfranchised almost entirely. Both the opposition and the right had denounced



the Bulygin Duma—the latter because it viewed the Bulygin Duma as a stepping stone to a dangerous Duma with legislative powers, the former because the Bulygin Duma fell far short of a legislative Duma based on four-tailed (universal, equal, direct, and secret) male suffrage. While a few ministers looked upon the idea of universal suffrage positively, the tsar and the majority of his advisors did not. After a series of meetings in October, the government approved an electoral law that was heavily influenced by that proposed earlier by Bulygin, yet it allowed a wider franchise. The complicated electoral law that was promulgated on December 11, 1905, while greatly increasing the number of voters, was not universal or equal or direct: it excluded women and accorded eligibility to those owning property or paying taxes (Ascher 1988, pp. 178–79; 1992, pp. 42–43, 79; Healy 1976, pp. 90–94, 100–104; Harcave 1964, pp. 246–47). Yet the government's efforts to create a conservative Duma bore results far from expected; in spite of the obstacles, the electorate voted for a Duma that was far more radical than imagined, with the Kadets the clear winners. Out of a total of 524 elected members, only 5% belonged to the ultraconservatives (Ascher 1992, pp. 50–52; Kochan 1966, p. 109).

Another legal attempt to reduce the Duma's powers was the decree issued on February 20, 1906, that transformed the State Council, a purely advisory body established in 1810, to a second chamber with legislative powers equal to those of the Duma. In a clear breach of the October Manifesto, which had not mentioned any legislative body other than the Duma, the newly transformed State Council (later incorporated into the Fundamental Laws) assured that no bill would be sent to the tsar for ratification unless it was approved by both houses. To guard the State Council's conservatism, the tsar not only appointed half of its 198 members, but he did so on a yearly basis, allowing quick removal of dissenters. The bulk of the rest of its members were elected from the established institutions such as the Orthodox church, and the provincial zemstvo assemblies, as well as the major landowners not represented in the zemstvos and the nobility or gentry. The State Council was thus transformed into a powerful conservative institution to greatly moderate, or even hamper, the Duma's attempts at social transformation and to act as a buffer between the tsar and the Duma (Ascher 1992, pp. 59–60, 70; Harcave 1964, pp. 13–14, 246–47; Healy 1976, pp. 109–10; Hosking 1973, pp. 11–12; Manning 1982, 212–13).

The Fundamental Laws, written entirely by the government and published on April 24, 1906, three days before the Duma's first session, exposed the Russian opposition's lack of input in the constitution-making process—further proof of their weakness. Given a free hand to draft the Fundamental Laws, the senior officials, in addition to the State Council,

incorporated many conservative features into the constitution. The Fundamental Laws preserved the powers of the executive branch with the tsar remaining as its absolute commander. Although the Duma achieved the right to interpellate ministers, the right to appoint or dismiss them was solely the tsar's and he could do so without the Duma's approval. Thus, the ministers remained responsible to the tsar rather than the Duma. Furthermore, the tsar had total command over the armed forces, summoned the sessions, and had the right to dissolve the Duma at will simply by indicating the election and convocation date of a new Duma. For a legislative proposal to become law, it had to pass both houses and gain the tsar's approval. In addition, the tsar reserved the right to veto all legislation, to determine foreign policy, to accept the views of either house in case of a budgetary disagreement, and—contrary to both houses' views—to retain the previous year's budget. He also had the right to issue emergency laws when houses were not in session. Furthermore, only the tsar could initiate revision of the Fundamental Laws. The only change in this period was a slight curbing of the tsar's power in the area of legislature (Ascher 1992, pp. 63–71; Harcave 1964, pp. 246–50; Healy 1976, pp. 118–20). The announcement came during the April congress of the Kadets, prompting their moderate leaders to make a severe condemnation of the government and the Fundamental Laws. "Like thieves in the dead of night, all the specialists on state law organized, [and] these people staged a conspiracy against the people (applause). . . . That which we read in the newspapers today is a fraud, a fraud against the people, and we must immediately answer this fraud," protested Miliukov, the Kadet leader (Ascher 1992, p. 79). On the same day of April 24, 1906, the government adopted its last precautionary measure against the yet unconvoked Duma by appointing the hostile Prime Minister Goremykin and a few archconservatives to his cabinet (Ascher 1992, pp. 73–76; Healy 1976, pp. 142–48).

During the short 72 days of its existence (April 27–July 9, 1906), the Duma endured a contentious constitutional battle with the government. The representatives' first action was to ask for sweeping changes designed to transform the Duma into a full-fledged parliament with absolute authority over the executive branch, such as that prevailing in the British constitutional monarchy. Among the representatives' requests submitted on May 2, 1906, in their "Answer to the Throne" were a cabinet responsible to the Duma, changes in the authority of or the elimination of the State Council, guarantee of civil liberties, the institution of universal male suffrage, the abolition of capital punishment, and amnesty for all political prisoners. Included was also the sensitive demand for agrarian reform, a topic discussed further below. Many of these demands violated the Fundamental Laws, and the government in its rejection of them

pointed this out. In response, on May 13 the Duma deputies cast a vote of no confidence by an overwhelming majority and asked for the resignation of the cabinet. The next day, the cabinet decided to dissolve the Duma, a decision that was withheld for some eight weeks. In the stormy days of the First Duma, the government approved only two of the Duma's legislative proposals and implemented a single one, while the representatives interpellated the ministers 400 times regarding their illegal actions and abuse of power (Ascher 1992, pp. 81–110, 162–71; Healy 1976, pp. 179–220).

What finally brought the Duma and the government to an impasse was the land question, the most pressing problem of the empire. In the Answer to the Throne, the Duma proposed alleviating the problems of the land-hungry peasantry by distributing treasury, monastic, and imperial lands, as well as through the "compulsory confiscation of private estates." These were followed by three particular proposals of various parties for solving the problem of the land-hungry peasants, all of which included compulsory confiscation of private land of major owners as an essential component. The government found compulsory expropriation of privately owned land "absolutely inadmissible" and instead proposed other solutions for raising agricultural productivity to improve peasants' conditions (Healy 1976, pp. 185–91, 221–38; Ascher 1992, pp. 171–77; Manning 1982, pp. 205–8).

The proposed reforms of the Russian Duma played the same role as they had in the Ottoman Empire and in Iran: they caused the breakdown of the previously constitutionalist coalitions and threw many among their ranks into the counterrevolutionary camp. Nowhere was this demonstrated more clearly than in the reaction of the landowning nobility. Greatly alarmed by the Duma's stance on the land question and by peasant unrest in the countryside, a majority of the landowning gentry turned sharply against the Duma and sided with the government that it had previously opposed.

The most dramatic and consequential of the coalition breakdowns was that of the landowning gentry that found its political preserve in the *zemstvos*, the local elective institutions of self-government established in the wake of Emancipation to fill the void left by the gentry as serf owners. Classless and representative in intent, these provincial institutions were, in reality, dominated by the landowning gentry. The landowning gentry contained a small emerging liberal minority that, with its urban residence, better education, professional orientation, and close ties to the nonnoble urban intelligentsia, constituted a faction distinguishable from the provincial gentry. Significantly, this liberal minority of the gentry intelligentsia occupied the leadership of the *zemstvos* and became an important advocate of the constitutional movement, joining the liberal Kadet party in large numbers. The *zemstvo* leadership was able to rally

zemstvo rank and file and other institutions of nobility behind the constitutionalist movement by exploiting their animosity toward the central government. This animosity was fueled by the government's neglect of agriculture in favor of its extensive industrialization policies and its intent to encroach upon the autonomy of zemstvos and turn them into instruments of central administration (Manning 1979, pp. 32–37; 1982, pp. 38–39, 43–49, 61–62; Verner 1990, pp. 105–6; Yaney 1973, pp. 235–38). In the words of Manning (1982, p. 49), “a cadre of gentry leaders within the zemstvos and other local elective institutions were able to translate the vague, often anti-bureaucratic sentiment of the provincial gentry into concrete political programs and demands.”

In a series of zemstvo congresses, first in November 1904 and then in April and autumn of 1905, the zemstvo leadership, along with broad ranging reforms, asked for a national legislative assembly, four-tailed male suffrage, and compulsory expropriation of private landholdings to alleviate land hunger among the peasantry. The provincial zemstvos followed the lead of congresses and became increasingly radicalized up to the winter of 1905–6. Yet at this time, notwithstanding its recent liberal voting, the provincial gentry became greatly alarmed by the recent peasant disorders and reacted by changing political position almost overnight. In face of government suppression of these revolts, they adopted silence, purged the left-wing “third element” zemstvo employees, and increasingly came to view their left-leaning Kadet leadership—with its program for expropriation of land—as traitors to the noble estate and unrepresentative of the local zemstvos (Manning 1979; 1982, pp. 67–137, 177–202; Harcave 1964, pp. 54–58, 132, 142–43, 171; Ascher 1988, pp. 60–65, 115–16).

When the First Duma convened, the debates on the agrarian reform sparked the second great wave of peasant uprisings in mid-May, during which close to 1,600 instances of peasant unrest were reported (Ascher 1992, pp. 111–28; Manning 1982, pp. 229–59). In face of the second wave of peasant disorders and the Duma's demand for compulsory expropriation, the formal split between the zemstvo leadership and the provincial gentry, or zemstvo rank and file, became solidified. Subsequently, the provincial gentry reacted by organizing congresses that mobilized the local zemstvos and noble societies behind a conservative political program, established the conservative United Nobility, and, in its 1906–7 elections, ousted the liberal leadership (Manning 1979; 1982, pp. 212–28).

Drawing its members from various conservative political landowning groups, the United Nobility was established during the Duma's agrarian debates as an alliance between the absentee magnates from St. Petersburg and the larger provincial landowners from regions hardest hit by peasant rebellions. A staunchly conservative institution, it acted as a counter-

weight to the power of the Duma by mobilizing the local zemstvos and noble associations, by exerting influence through legal channels of the State Council, and by its behind-the-scenes influence on the tsar and senior officials. It fiercely opposed compulsory expropriation of the gentry's lands and instead suggested abolition of the commune as the solution to the land problem, a program later adopted by the government. Furthermore, its political machinations played an influential role in convincing the government to abolish the First Duma. Until its demise in 1915, it increased its presence in formal-legal institutions, such as the State Council, and continued to influence the government through informal machinations, steering its policy in a more conservative direction (Hosking and Manning 1979; Manning 1982, pp. 229–59; Ascher 1992, pp. 4–5, 178–80, 196; Harcave 1964, pp. 256–57; Levin 1966, pp. 237–38).

The ease with which the government moved to abolish the First Duma was a clear indication of the Duma's lack of power; despite the various appeals the Duma issued, it failed to mobilize popular support or to incite widespread disturbances in opposition to the government's decision. In Iran and in the Ottoman Empire, the assemblies were abrogated forcefully after armed confrontations with their supporters, for they constituted genuine power blocs against the government. In Russia the Duma was annulled by decree without causing any major disturbances.

Deadlocked on the question of expropriation of land, the Duma issued its famous "Appeal to the People," indirectly threatening the government with future public disturbances if its proposed land program was not approved. The appeal prompted the government to issue a decree and to dissolve the Duma on July 9, 1906. In response, in an emergency meeting that convened outside St. Petersburg on the same day, the deputies issued the Vyborg Manifesto, calling the public to massive revolutionary disturbances in response to the dissolution. Even though the revolutionary manifesto was eventually signed by a surprisingly large number of deputies, it failed to incite any large-scale response on the part of the workers or the peasants, a situation that was not unexpected; the Duma had not indulged in active organization of the opposition at the time of its operation, while the public, and in particular the workers, had not only received severe blows to their political associations but after a year and a half of revolutionary struggle had exhausted their resources and could not bear the threat of unemployment (Ascher 1988, pp. 192–209; Healy 1976, pp. 238–61; Manning 1982, pp. 241–43, 260–61).

With the dissolution of the First Duma, the government set on the task of further limiting the franchise to create an "acceptable" Duma. With this intent, "all the resources of the bureaucracy, physical, legal, and spiritual, were brought into play in an attempt to create a duma

with a majority to the right of the center, a duma 'acceptable' to the administration" (Levin 1966, p. 60). Under the new electoral law, as the result of combined class, property, and territory requirements, it became possible for a single individual to vote several times (Levin 1966, pp. 60–62). But even this electoral law failed the bureaucracy and only gave birth to a more radical Duma harboring greater hostility toward the government, though with a larger concentration of the extreme right-wing faction (Levin 1966, pp. 60–69; Ascher 1992, pp. 284–85; Kochan 1966, p. 117).

During the 103 days of its existence and its 53 meetings between February 20 and June 1907, the Second Duma's legislative activities did not fare any better than the First's. The Duma continued to press the government with its interpellations, asking for responsible ministries and questioning the government's illegal activities, while it entered major disagreements with the government over the budget. Yet, the agrarian issue, as with the First Duma, continued to be the most contentious issue, and it finally brought the Second Duma's demise as well. In response to the Duma's program of extensive compulsory expropriation of land, the government reacted with a coup d'état that dissolved the Duma and promised new electoral laws. With this act, the government set out to regain what it had lost to the opposition during the last two years, violating the principles stated in the October Manifesto and the constitution on its way to full restoration of its own powers (Levin 1966, pp. 156, 186, 198–99, 202, 222, 227, 242–60, 307–8; Ascher 1992, pp. 318–20, 349–58; Manning 1982, pp. 328–29; Bonnell 1983, pp. 319–21). It was only during the Third Duma that the government, with its modified electoral laws, could give the majority representation to the landowners and drastically decrease the number of workers' and peasants' representatives inside the chamber (Levin 1966, pp. 340–41; Manning 1982, pp. 325–30, 357–59; Ascher 1992, pp. 353–55; Hosking 1973). Nevertheless, even this moderated Duma was not deemed acceptable to the conservative bureaucracy.

It would be an exaggeration to conclude that the Russian bureaucracy, in its entirety, resisted all change. Most notably, the prime ministers Witte and Stolypin attempted to reach some form of accord with the opposition. Yet, the predicament of the reform-minded bureaucrats was that they were part of an administrative organ that was overwhelmingly dominated by conservatives who frustrated their efforts at change. As Ascher has pointed out, "Devotees of the old order held the vast majority of posts in all branches of the bureaucracy" (Ascher 1988, pp. 245–48, quote from 245; Harcave 1964, p. 245; Healy 1976, p. 134). The conservatives, aside from their legal interventions, did not hesitate to use illegal means to achieve desired results. During the elections for the First Duma

they cooperated with both the rightist gangs, who attacked the revolutionaries, and the police, who frequently shut down election meetings (Ascher 1992, pp. 44–45; Healy 1976, p. 127). These illegal activities continued during the Second and Third Dumas. As Levin in his investigation of the Second Duma has noted, anti-Duma activities were not confined to the highest ranks of the bureaucracy, but ranged from the governors and city police chiefs to the lower ranks of bureaucracy and local officials responsible to the center (Levin 1966, p. 64). Notable was the moral and financial support the administration granted the extreme-right Union of Russian People, an act that enabled the union to wield an influence far greater than its actual numbers.<sup>76</sup> During the Third Duma, it was the bureaucracy's resistance to even the slightest change that led St. Petersburg's high-ranking chief of police and other cabinet ministers to assassinate Prime Minister Stolypin (Hosking 1973, pp. 148–49).

Unlike the landowning gentry, who only gradually came to realize the threat of the constitutional movement, the church was opposed to the October Manifesto from the very beginning and remained so by backing the government during its election campaigns and resisting the Duma's proposed reforms (Ascher 1988, p. 244; Levin 1966, p. 64). On the other hand, at least at the beginning, some industrialists tenaciously supported the constitutional movement; a substantial number of the business class had participated in the general strike that led to the October Manifesto. Yet, even then, their position was not unanimously favorable, and many industrialists demanded stern measures against the labor strikes. The opinion of the opposing faction prevailed in the aftermath of the general strike and the growing labor militancy during the Days of Freedom. On October 24, 1905, the Association of Manufacturers and Factory Owners, representing 150 companies, formed to implement a uniform strike policy among the employers. With the legalization of trade union activity after the March 1906 laws, employers were also presented with a more favorable legal framework to organize in various societies and associations to defend their interests and oppose the demands of labor and the reforms of the Duma (Ascher 1988, pp. 121–22, 216–17, 222, 279–80; Bonnell 1983, pp. 274–318; Hosking 1973, pp. 179–80).

### Resistance across Constitutional Revolutions

To sum up, the dynamics of counterrevolution were broadly the same in all three revolutions. The highly destabilizing programs of the assemblies

<sup>76</sup> The tsar supported their newspaper financially from 1906 to 1909 (Levin 1966, pp. 238–39).

were initiated at a time when the revolutionaries lacked sufficient state control. The assemblies' undertakings not only disturbed the old regimes and their bureaucrats, they managed to antagonize even some of the social sectors that had avidly supported the constitutional system. The old regimes, still largely in state command, allied with the negatively affected groups to overturn the tide of revolution and force the assemblies out of power. In Russia, the Duma was first weakened by the violent destruction of a potential ally—the soviets—and then it was gradually weakened through a series of legal, constitutional procedures until it was left without any real powers. While the Russian Duma never recovered from the counterrevolutionary backlash, the Ottoman and Iranian constitutionalists did. The following section is devoted to explaining this difference.

#### DEFEATING THE COUNTERREVOLUTION

After being forced out of power, the Ottoman and Iranian revolutionaries were able to defeat the counterrevolution and restore the constitutional regimes. The support of a major sector of the armed forces proved to be crucial in this task. The role of the military, the most crucial factor for success, has been the subject of extensive commentary in theories of revolution from a variety of perspectives (Brinton 1952, pp. 98, 280; Pettie 1938, pp. 102–106; Tilly 1978, p. 200; Skocpol 1979, p. 32; Russel 1974, pp. 9, 79–82; Gurr 1970, pp. 244–48, 251–56, 272; Stone 1966, p. 166; Eckstein 1965, p. 157). However, many have discussed the military in terms of its ability or inability to repress the contenders rather than as a potential and indeed crucial ally of the opposition. Only a few have taken the further step of paying attention to the fact that if the opposition is to succeed, it should command at least a faction of the armed forces (Brinton 1952, pp. 98, 280; Russel 1974, pp. 9, 79–82). Brinton has remarked (p. 98) that “no government has fallen before revolutionists until it has lost control over its armed forces or lost the ability to use them effectively; and conversely that no revolutionists have ever succeeded until they have got a predominance of effective armed forces on their side.”

Second in importance were the quasi-governmental institutions established and further consolidated when the revolutionaries' activities were legally sanctioned. The presence of such support was decisive in dealing with the counterrevolution when it was gathering momentum and once it was in full swing. Naturally, the role of such institutions was most crucial in Iran and Russia, where the constitutionalists had the least amount of support from within the state. Yet, the Iranians fared better



because the committees, in the period of legal activity, gathered sufficient momentum to be able to continue their operations even with the onset of the counterrevolution. In Russia, the soviets were crushed and banned and the zemstvos turned against the constitution, thus leaving the Duma and the constitutionalists without support to stave off or to recover from the counterrevolution.

### Restoring Power in the Ottoman Empire

After the CUP was unexpectedly forced out of power on April 13, 1909, it retreated to its stronghold in Salonika. From there, within a mere 10 days, it organized the Action Army (*Hareket Ordusu*) and easily captured Istanbul.<sup>77</sup> Subsequently, the Chamber of Deputies was restored and its unionist members reappointed. Furthermore, after the CUP accused Abdülhamid II of leading the movement, he was deposed and replaced with his brother Mehmet V.<sup>78</sup> Some 200 participants, predominantly soldiers but also members of religious societies such as the leader of the Society of Muhammad, were publicly hanged, and 10,000 soldiers were punished, many of them by banishment to the CUP-dominated provinces as simple laborers for public works projects.<sup>79</sup>

<sup>77</sup> See Nadi (1909, pp. 145–51). Some historians, by relying extensively on the rhetoric of the CUP, have underplayed the CUP's connection to the Action Army. My claim is that such disassociation was a later development, motivated by political expediency on the part of the CUP. Attempts at distancing, e.g., may be found in the statements of Mahmud Şevket Pasha, the commander of the Action Army (Danışmend, 1961, pp. 134–35). Likewise, when the Action Army captured Istanbul, strikingly, their public announcements did not make any mention of the CUP (see *İkdam*, no. 5360, April 29, 1909 (9 Rebiyülâhır 1327), p. 1; no. 5363, May 2, 1909 (12 Rebiyülâhır 1327), p. 1. It seems certain however, that the Action Army was organized by the CUP because it originated in Macedonia, the center for the CUP activities and the original cite of mutinies and disorders. In addition, the CUP had earlier telegraphed the sultan, grand vizier, and other high-ranking officials with a warning: CUP was organizing an army to retake power and to restore the chamber with Ahmed Rıza, the former head of the chamber and a highly prominent CUP member, as its president (Danışmend 1961, pp. 99–108). In light of the above it is reasonable to agree with Danışmend that Mahmud Şevket Pasha, the commander of the Action Army, was appointed by the unionists.

<sup>78</sup> An announcement by the Action Army condemned Abdülhamid II for the entire counterrevolutionary incident; see *İkdam*, no. 5360, April 30, 1909 (9 Rebiyülâhır 1327), p. 1. In contrast to the public declarations of the CUP, the *fetva* issued by *şeyhülislam* mentioned other matters as the main reasons for dethronement; *Düstur*2, I, no. 57, April 27, 1909 (7 Rebiyülâhır 1327), p. 166.

<sup>79</sup> See *BDFA* (1910, p. 112), Tunaya (1984, p. 190), and *Düstur*2, I, no. 64, May 30, 1909 (10 Cemaziyelevvel 1327), p. 191.

### Defeating the Monarchy in Iran

The Iranian revolutionaries did not recover as easily and remained out of power from June 23, 1908, to July 16, 1909. By the time the National Assembly was sacked in Tehran, the committees had built strong organizations in the provinces. This was especially true in the north and northwest. Tabriz, the city where the committees were strongest, resisted the government's onslaught and remained a major actor in the national resistance movement. The reinstitution of the constitutional government, however, was made possible only after the Bakhtiari tribes in the south joined forces with the committees. Persuaded by the Isfahan committees in early January 1909 to join the national resistance movement, the Bakhtiaris took possession of Isfahan and deposed the government representatives. This was followed by a series of victories for the committees in the north, including the takeover of Rasht and the defeat of government forces in Tabriz. The joint cooperation of the committees—predominantly from the north with a faction of the Iranian armed forces (the Bakhtiaris) from the south—resulted in a four-day battle for Tehran, which ended on July 16, 1909. The shah, who had played a central role in the counterrevolution, was deposed on the same day and two days later his young son, a minor, was proclaimed shah and placed under the supervision of a regent. The National Assembly no longer had to contend with the government of the old regime (Browne 1910, pp. 266–327; Kasravi 1951, pp. 640–906; Lambton 1963, pp. 67–76).

As the discussion on prerevolutionary reforms pointed out earlier, the Iranian army was rife with major structural divisions, the most notable of which was that between the standing army and the tribal factions. It was precisely the existence of this structural division that proved crucial during the revolutionary upheavals. The tribal faction's decision to collaborate with the constitutionalists was both in reaction to the old regime's reform policy of weakening the tribal forces and a response to an opportunity to bolster their own power by cooperating with a challenger to the centralizing state, a path well trodden by nomadic tribes throughout Iranian history. The revolutionaries exploited this structural cleavage within the military to their own advantage when they persuaded the Bakhtiaris to unite with them.

### The Apathetic Army in Russia

In Russia, despite the army's low pay and the dismal state of its rank and file, it remained apathetic toward the revolutionaries.<sup>80</sup> While a state

<sup>80</sup> For an extensive treatment of the army during the Revolution of 1905 see Bushnell (1985); see also Ascher (1988, p. 168).

financial crisis has never been sufficient to throw an army into the revolutionary camp, the losing war that Russia continued to fight against Japan could have placed Russia's military on the side of revolution. A crucial contingency that kept the Russian army from developing deeper and more fundamental grievances and ultimately saved the autocracy from soldiers' defection was that the Russian statesmen, alarmed by the internal disturbances, concluded the war with Japan quickly, before the soldiers at the front were affected by the disturbances at home. Domestic conditions forced the Russian leaders to accept defeat at the hands of Japan, humiliating as it was, and concentrate their energy and resources on the home front without further antagonizing the defeated army. If the revolutionaries had been able to win the cooperation of a major sector of the army and navy, they could have used the military as an essential resource to orchestrate a full-fledged revolution (Ascher 1988, pp. 167–68).

There were many scattered mutinies in the army, but few were connected to the urban revolts. The soldiers and sailors did not identify with the general demands of the movement and failed to join workers and revolutionaries in the uprisings against the autocracy. When they did revolt, it was inward looking: they demanded only the improvement of their conditions. In short, with rare exceptions, such as the mutiny on the battleship *Potemkin* and at the Kronstadt naval base, the soldiers' uprisings did not echo the broader political demands of the movement. The Social Democrats had made only a marginal penetration into their ranks, but even at the front, the soldiers remained unaffected by the revolutionary propaganda. The elite forces, and the Cossacks in particular—the most effective instrument in fighting the revolution—were untouched by the general revolutionary fervor and remained absolutely loyal to the tsar (Ascher 1988, pp. 170–74, 269–73, 311–12, 325; Bushnell 1985, pp. 226–28; Harcave 1964, pp. 42–43, 140–41, 156–57, 220–22; Healy 1976, p. 70).

Dire financial conditions and a losing war were strong enough negative stimuli to make the Russian military an unreliable instrument of repression and an unpredictable force for both the autocracy and the revolutionaries. It was precisely this unpredictability after Bloody Sunday but before the granting of the October Manifesto that had forced the autocracy to accept the demands of the opposition (Ascher 1988, p. 168; Verner 1990, pp. 161–62; Bushnell 1985). Yet, after the announcement of the October Manifesto, the military proved to be a reliable ally of the autocracy at the onset of the serious outbreaks of labor and peasant unrest (Ascher 1988, p. 168). Prompt acceptance of the defeat against Japan, proved to be the right course of action for the Russian statesmen, but after entering the First World War, the timely conclusion of the war was a luxury they could no longer afford.

### Advancing to the Fourth Stage

The argument presented so far may be summarized in table 1. During stage 2, the assemblies' effectiveness in approving legislation, implementing legislation, dismissing and replacing individual ministers, prime ministers, or entire cabinets with sympathizers, and finally, modifying the constitution or interpreting it to their own advantage all depended upon the support they received from the revolutionary power bloc formed in their support. The most important elements of these power blocs were the military officers and civilian bureaucrats; next in importance were the quasi-governmental institutions and their civilian militia. It should be noted that the second stage of constitutional revolutions provides a crucial period of legal activity for the contenders, enabling them to increase their influence within the state administration and the armed forces and to expand considerably their numbers, their scope of activity, and the powers of quasi-governmental institutions and their militia.

With the exception of organizing a militia, these activities were successfully pursued by the Young Turks during stage 2, a fact that explains the success of the Ottoman Chamber of Deputies during its constitutional battles with the government and the ease with which the chamber defeated the counterrevolution within 10 days. The extensive support the CUP enjoyed within the armed forces had rendered the formation of a militia unnecessary. The period of legal activity proved to be most crucial for the Iranian revolutionaries. The Iranian constitutionalists lacked support from either the military or the bureaucracy and enjoyed only the backing of the committees and the militia they organized during the legal period. With the help of these institutions, the constitutionalists were able to intimidate the government and to withhold the abrogation of the constitutional system and the destruction of the assembly even in the face of grave legal violations for more than 20 months. They even empowered the assembly to the degree that it could win a few legal battles against the government. On the other hand, after the onset of the counterrevolution and the fall of the National Assembly, the committees sustained the revolutionary movement in the provinces for more than a year until a major faction of the military forces decided to assist them in defeating the counterrevolution and restoring the assembly. The Russian Duma, on the other hand, was not an effective institution because it did not have access to any extra-parliamentary sources of support: the soviets were suppressed early, the zemstvos rescinded their support, and it had hardly any allies within the bureaucracy or among the mutinous soldiers. The constitutionalists were therefore unable to forestall the legal counterrevolution that, piece by piece, changed the Duma's representation and abolished its powers.

TABLE 1  
OUTLINE OF CONSTITUTIONAL REVOLUTIONS

ENTITY	SUPPORT FROM INSIDE THE STATE			QUASI-GOVERNMENTAL INSTITUTIONS ESTABLISHED BY THE REVOLUTIONARIES			RESULTS	
	Bureaucracy	Army		Institutions with Civil Duties	Militia		Assembly's Performance and Power	Defeated Counter- revolution
Ottoman Empire	Yes	Yes	(extensive)	Yes	NA*		Impressive	Yes
Iran	No	Yes	(no initially but yes finally)	Yes	Yes		Moderate	Yes
Russia	No	No		No	No	(yes initially but no finally)	Poor	No

\*The CUP had a strong presence within the military and did not see a need to create paramilitary organizations by arming its clubs and branches

## DISCUSSION: STATE BREAKDOWN

At the beginning of this article I explained that the modernizing Ottoman reforms that began in the late 18th century and lasted until the revolution in the early 20th century created pronounced divisions within the state. Furthermore, throughout this discussion I have held that the Ottoman bureaucracy and military became revolutionary because of internal divisions within the state and the relative deprivation that the modern-trained bureaucrats and military officers experienced against the less educated, traditionally educated, military and civil bureaucrats who had patrimonial connections to the *grandees* or to the sultan himself. Despite my consistent emphasis on the role of resources and resource mobilization theory, these assertions may appear more familiar to theories of relative deprivation and at odds with resource mobilization theory, which attributes the outbreak of revolts and revolutions not to a change in the actors' grievances, but to the type and amount of resources available to actors.<sup>81</sup> Yet, it becomes much harder to uphold resource mobilization's claim when it comes to the analysis of resourceful actors—actors who have routine and regular access to resources, such as members of the civil bureaucracy and military (members of "polity" in Tilly's [1978] classification). For this set of actors, the level of access to resources is constant; the variability of their actions may be better explained by the level of variability in their grievances.

The above explanation should by no means be taken as a rejection of resource mobilization or state-centered theories, according to which revolutionary outbreaks occur only when states are incapable of performing their routine tasks (Brinton 1952, pp. 30–32, 37, 41, 279–80; Pettee 1938, pp. 100–101; Skocpol 1979, p. 32; Goldstone 1991b). As George Pettee remarked, "[Revolution begins] simply with a sudden recognition by almost all the active and passive membership that the state no longer exists" (p. 100). A weaker state has a lower capacity to suppress the challengers, a condition that translates into lower costs of mobilization for the contenders. In agreement with state-centered theories, I found that three states that faced revolutionary situations were financially troubled (Skocpol 1979; Goldstone 1991b).

Yet, fiscal crisis alone, without the existence of internal divisions

<sup>81</sup> Resource mobilization theory has certainly done away with some simplified assumptions of the relative deprivation theory and has convincingly shown that, no matter how aggrieved, without the existence of resources, revolutionary actors are not able to protest and contest authority. For a view opposed to relative deprivation theory see Tilly, Tilly, and Tilly (1975) and McCarthy and Zald (1977). For one of the most convincing empirical demonstrations of this claim see Aminzade (1984). For the classic presentation of relative deprivation theory see Davis (1962, esp. p. 6). See also Gurr (1970, pp. 22–154).

within the civil and military administrations, would not lead to state breakdowns. After all, often when a state is fiscally challenged, the pay for the army and the bureaucracy is in arrears, a situation that particularly affects the lower ranks. This source of grievance, however, has never been sufficient to prompt internal actors to side against the old regimes. Instead, the strongest impetus for dissolving the bonds of loyalty with the old regime comes from pronounced structural divisions within the state. A good counterexample may be found in the reaction of the Russian army to the revolutionary outbreaks. Because of the modernizing reforms that had created administrations devoid of pronounced structural divisions, the Russian army, despite its dire conditions, refused to join the ranks of revolutionaries. Although it rebelled against its own destitute conditions, its actions remained internal and did not compromise its loyalty toward the old regime. Thus, I hold that the analysis of the state structures before the revolution may lead the way for constructing a more precise definition of state breakdown, a definition that could differentiate between (1) mere internationally challenged and financially troubled states, (2) states that, because of fiscal crisis, have lost control over large territories and are incapable of enforcing the rule of law, and (3) states that, in addition to all of the above, witness the cooperation of large sectors of their military and bureaucracy with revolutionaries. If the military and bureaucracy are torn by structural divisions, they may, at times of financial and international difficulties, readily break off from the state and side with the opposition in hope of future gains from other sectors. Greater attention to the internal structuring of states may serve as a corrective to more recent theories that, unlike earlier theories' emphasis on the bureaucratic and military staff (Brinton 1952; Pettee 1938), have descended into a more abstract and less precise definition of breakdown.

Early 20th-century Iran is another instance in which internal state divisions came to the constitutionalists' rescue. As discussed above, despite the attempts of Qajar reformers, the Iranian army continued to include important semiautonomous divisions of tribal cavalry. A faction of these tribal cavalries joined the constitutional militia to end the counterrevolutionary backlash and restore the constitutional regime. These divisions had repeatedly been put to use by rival claimants to the throne, and now the constitutionalists exploited this division with the promise of a greater future role in politics for the cooperating tribal faction.

Although I have claimed that prerevolutionary divisions within the state are the most probable cause of state breakdowns, this claim does not rule out other possible conjunctural causes for breakdown. Chief among these conjunctural causes is severe defeat in a long war, prompting the soldiers to blame their governments for the suffered rout (Gurr

1970, p. 254; Stone 1966, p. 166). Considering the financial crisis of the state and widespread disturbances as a constant, my claim is that soldiers defect and states breakdown sooner if there are already preexisting structural divisions within administrations and armies. The structural divisions that prevailed in the Ottoman state and in the Iranian army acted as a potent impetus for dissolving the loyalty of their staffs, without the necessity of immediate international setbacks.

## CONCLUSION

The radical Iranian newspaper *Musavat* (Equality) in the heat of the revolutionary struggles stated, "One cannot read a single page of a newspaper without coming across the word liberty at least ten times and one cannot walk twenty steps in the alleys of Tehran without coming across one or two committee banners, adorned in bold characters with the three holy phrases of liberty, fraternity and equality."<sup>82</sup> On the last day of the soldiers' mutiny in the western regions of the Ottoman Empire, the British Vice Consul reported a typical instance of the proclamation of the constitution by a staff officer addressing soldiers and officers under his command. In his speech "he laid great stress on the absolute necessity of establishing a constitutional Government if the Turkish Empire was to be saved from utter ruin, on the abolition of distinction between race and creed under the Ottoman Government and gave the words 'la patrie, liberté, égalité, fraternité' as their motto" (as quoted in Ahmed 1969, p. 12). In the aftermath of the Young Turk Revolution and in the heat of the day-to-day power struggles with the old regime, *Tanin*, the official organ of the CUP, published regular columns on the history and analysis of the great French Revolution, a conflict whose lessons were thought to be a guiding light for the current situation. The profusion of translations and analyses of the history of the French Revolution in the postrevolutionary setting was a grave contrast to the situation under Hamidian censorship, during which a passing reference to the "regime of 1789" in a journal brought its suppression for a few weeks in 1901. (See Lewis 1961, pp. 192-193; see also pp. 195-96. For the earlier impact of the French Revolution see Lewis [1953, pp. 105-25].) The relevance of the French Revolution was nowhere sensed more clearly than in Russia in 1905, where liberals, moderate radicals, the Social Democrats, and the anarchists drew on it for their own uses, while the monarchists also exhibited an intense interest in French history and reacted to the opposi-

<sup>82</sup> *Musavat*, no. 25, May 31, 1908 (29 Rabi' II, 1326), p. 7. For a description and analysis of the French Revolution see *Musavat*, no. 7, December 3, 1907 (27 Shawwal 1325), p. 1.



tion having in mind the fate of the aristocracy under Louis XVI (Shlapentokh 1988, pp. 254–401; Verner 1990, pp. 116, 219–20; Keep 1968). And unlike Iran and the Ottoman Empire, where the lessons of the French Revolution seemed mostly to interest the young intellectuals, in Russia the singing of “*La Marseillaise*” by the ordinary participants during various instances of the conflict throughout the empire indicated greater familiarity at a popular level (Ascher 1992, pp. 54, 133, 158, 208–9; 1988, p. 230; Harcave 1964, pp. 199–200; Bonnell 1983, p. 168; Shlapentokh 1988, p. 257). According to Shlapentokh (1988, p. 257), in Russia between 1905 and 1917, “the French Revolution became more popular than any other period of world history, even including Russian history.”

Reference could be made to many other instances to prove that the French Revolution loomed large, at least in the minds of some participants. The contribution of the French Revolution, however, was more subtle than the above may suggest, for it is impossible to imagine that all actors were aware of the French Revolution and that they attempted to emulate that revolution at every step. Historical evidence does not support such an argument. Nor was the revolutionaries’ major demand, a call for a constitutional system, seen to be peculiar to France alone. By the early 20th century, constitutionalism had taken a strong hold in all of Western Europe, and the constitutionalists idealized the system of rule throughout Europe rather than the one particular to France. Yet, in all three settings, France occupied a privileged position in the constitutionalist consciousness, because France was deemed to be the first constitutional model, the rest of Europe was considered to have followed in its footsteps, and in France constitutionalism was acquired through revolutionary means.

The crucial contribution of the French Revolution was that it made available the revolutionary paradigm of constitutionalism, a paradigm that structured the relationship of the challengers with the old regimes. The revolutionaries, instead of demanding the complete and sudden overthrow of the old regimes, asked for the creation of an assembly by means of which they intended to render the traditional structures of rule ineffective. This path to power gave the constitutional revolutions an altogether different dynamic than revolutions that took place after the Russian Revolution of 1917. By placing the revolutions on a distinctly different trajectory, the ideology of constitutionalism had a profound impact upon the manner in which the struggle for power unfolded.

As I have argued in this article, once the ideology of constitutionalism placed the revolutions on a similar trajectory, it was the support of extra-parliamentary resources, including the military, that determined (1) the level of success of the assemblies and (2) whether the revolutionaries could defeat the counterrevolution and advance to the final stage to cap-

ture total state power. It was during the second stage, the period of legal activity after the issuing of decrees that a crucial time factor entered into the constitutional revolutionary processes. This time factor allowed the revolutionaries to build quasi-governmental institutions and to penetrate deeper into the state. This was, it should be noted, a critical period that is missing from the socialist revolutionary processes. As I demonstrated above, the Young Turks used this opportunity to the greatest extent by reorganizing and purging officials of the old regime and by increasing their illicit party activities around the Ottoman Empire, the Iranians by building and legalizing committees, and the Russians by consolidating labor organizations. Yet, because the period of legal activity ended in Russia at an early stage, the Russian assembly enjoyed little success in its legislative activities.

For a more complete typology of revolutions, further research may take three major directions. First, comparisons to other constitutional revolutions, such as those in China and Mexico, should help confirm, reject, or qualify the conclusions reached here. Second, investigation of the processes of the socialist revolutions should bear out in greater detail their differences from constitutional revolutions. Last, research may greatly benefit by concentrating on what I would call transition paradigms. The 1789 French and 1917 Russian revolutions belong to this category. The French Revolution started out as a revolt that only at the end became a constitutional revolution (for a concise history see Lefebvre [1947]). Similarly, the 1917 Russian Revolution was heavily influenced by the French Revolution, started out as a constitutional revolution, and followed closely the dynamics of the 1905 revolution (for a brief history see Fitzpatrick [1982]). Yet, its difference in some crucial respects in its later periods set it apart from the previous generation of constitutional revolutions. Acting as a transition paradigm, it stood between the constitutional revolutions of the 19th and early 20th centuries and the later generation of socialist revolutions, which, in contrast to constitutional revolutions, demanded a sudden and violent overthrow of the old regimes from their inception. Future research should address the conditions under which these transitions take place.

## REFERENCES

- Abbott, Andrew. 1983. "Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes." *Historical Methods* 16, no. 4 (Fall): 129-47.
- . 1990. "Conceptions of Time and Events in Social Science Methods. Causal and Narrative Approaches." *Historical Methods* 23, no. 4 (Fall): 140-50.
- . 1991. "The Order of Professionalization. An Empirical Analysis." *Work and Occupations* 18, no. 4 (November): 355-84.

- . 1992. "What Do Cases Do?" In *What Is a Case? Exploring the Foundations of Social Inquiry*, edited by Charles C. Ragin and Howard S. Becker. Cambridge: Cambridge University Press.
- Abbott, G. F. 1909. *Turkey in Transition*. London: Edward Arnold.
- Adamiyat, Faridun. 1969. *Amir Kabir va Iran*. Tehran: Khvarazmi.
- . 1972. *Andishah-i Taraqqi va Hukumat-i Qanun: 'Asr-i Sepahsalar*. Tehran: Khvarazmi.
- . 1976. *Idi'uluzhi-i Nahzat-i Mashrutiyyat-i Iran*. Tehran: Intisharat-i Payam.
- Afshar, Iraj. 1989. *Qabalah-i Tarikh*. Tehran: Talayah.
- Ahmad, Feroz. 1969. *The Young Turks: The Committee of Union and Progress in Turkish Politics, 1908–1914*. Oxford: Clarendon.
- Akşin, Sina. 1971. *31 Mart Olayı*. Istanbul: Sinan Yayınları.
- Algar, Hamid. 1969. *Religion and State in Iran, 1785–1906: The Role of the Ulama in the Qajar Period*. Berkeley: University of California Press.
- Aminzade, Ronald. 1984. "Capitalist Industrialization and Patterns of Industrial Protest: A Comparative Urban Study of Nineteenth-Century France." *American Sociological Review* 49 (August): 437–53.
- Arendt, Hannah. 1977. *On Revolution*. New York: Penguin.
- Arjomand, Said Amir. 1981. "The Ulama's Traditionalist Opposition to Parliamentarism, 1907–1909." *Middle Eastern Studies* 17, no. 2 (April): 174–90.
- . 1988. *The Turban for the Crown: The Islamic Revolution in Iran*. New York: Oxford University Press.
- . 1992. "Constitutions and the Struggle for Political Order: A Study in the Modernization of Political Traditions." *European Journal of Sociology* 33:39–82.
- . 1993. "Religion and Constitutionalism in Western History and in Modern Iran and Pakistan." In *The Political Dimensions of Religion*, edited by S. A. Arjomand. Albany: State University of New York Press.
- Ascher, Abraham. 1988. *The Revolution of 1905: Russia in Disarray*. Stanford, Calif.: Stanford University Press.
- . 1992. *The Revolution of 1905: Authority Restored*. Stanford, Calif.: Stanford University Press.
- Bakhash, Shaul. 1978. *Iran: Monarchy, Bureaucracy and Reform under the Qajars, 1858–1896*. London: Ithaca Press.
- . 1983. "The Failure of Reforms: The Prime Ministership of Amin al-Dawla, 1897–1898." In *Qajar Iran: Political, Social and Cultural Change, 1800–1925*, edited by E. Bosworth and C. Hillenbrand. Edinburgh: University of Edinburgh Press.
- BDFA (*British Documents on Foreign Affairs*). 1908. Vol. 20, doc. 18, October 23.
- . 1909. Vol. 20, doc. 24, March 1.
- . 1910. Vol. 20, doc. 26, January 31.
- Bendix, Reinhard. 1984. *Force, Fate, and Freedom: On Historical Sociology*. Berkeley and Los Angeles: University of California Press.
- Berkes, Niyazi. 1964. *The Development of Secularism in Turkey*. Montreal: McGill University Press.
- Birge, John Kingsley. (1937) 1965. *The Bektashi Order of Dervishes*. London: Luzac.
- Bonnell, Victoria E. 1983. *Roots of Rebellion: Workers' Politics and Organizations in St. Petersburg and Moscow, 1900–1914*. Berkeley and Los Angeles: University of California Press.
- Braudel, Fernand. 1980. "History and the Social Sciences: The Longue Durée." Pp. 25–54 in *On History*. Chicago: University of Chicago Press.
- Brinton, Crane. (1938) 1952. *The Anatomy of Revolution*. New York: Prentice-Hall.
- Brown, John P. (1867) 1927. *The Dervishes*. London: Oxford University Press.
- Browne, Edward G. 1910. *The Persian Revolution of 1905–1909*. Cambridge: Cambridge University Press.

- Burns, Gene. 1991. "Instability and Ambiguity in Revolutionary Coalitions: The Case of Iran." Paper presented at the 1991 meeting of the American Sociological Association, Cincinnati.
- Burujini, Jamshid Zargham. 1971. *Dawlatha-yi 'Asr-i Mashrutiyyat*. Tehran. Idarah-i Kull-i Qavanin.
- Bushnell, John. 1985. *Mutiny amid Repression: Russian Soldiers in the Revolution of 1905-1906*. Bloomington: Indiana University Press.
- Buxton, Charles Roden. 1909. *Turkey in Revolution*. London: Fisher Unwin.
- Buxton, Noel. 1909. "The Young Turks." *Nineteenth Century and After* 65:16-24.
- Cevat Bey, Ali. 1960. *İkinci Meşrutiyetin İlanı ve Otuz Bir Mart Hadisesi*, edited by Faik Reşit Unat. Ankara: Türk Tarih Kurumu.
- Danışmend, İsmail Hami. 1961. *Sadr-i-a'zam Tevfik Paşanın Dosyasındaki Resmi ve Hususi Vesikalarla Gore: 31 Mart Vak'ası*. Istanbul: İstanbul Kitabevi.
- Dawlatabadi, Yahya. 1983. *Hayat-i Yahya*, 4 vols. Tehran: Intisharat-i 'Attar, Intisharat-i Firdawsi.
- Davis, James C. 1962. "Toward a Theory of Revolution." *American Sociological Review* 27, no. 1 (February): 5-19.
- Davison, Roderic. 1963. *Reform in the Ottoman Empire, 1856-1876*. Princeton, N.J.: Princeton University Press.
- Downing, Brian M. 1988. "Constitutionalism, Warfare, and Political Change in Early Modern Europe." *Theory and Society* 17, no. 1 (January): 7-56.
- Eckstein, Harry. 1965. "On Etiology of Internal Wars." *Theory and Society* 4: 133-63.
- Edwards, Lyford. (1927) 1970. *The Natural History of Revolution*. Chicago: University of Chicago Press.
- Fainsod, Merle. 1963. "Bureaucracy and Modernization: The Russian and Soviet Case." In *Bureaucracy*, edited by La Palombara. Princeton, N.J.: Princeton University Press.
- Farhi, David. 1971. "The Şariat as a Political Slogan—or the 'Incident of the 31 Mart.'" *Middle Eastern Studies* 7 (3): 275-99.
- Findley, Carter. 1980. *Bureaucratic Reform in the Ottoman Empire: The Sublime Porte, 1789-1922*. Princeton, N.J.: Princeton University Press.
- . 1989. *Ottoman Civil Officialdom, a Social History*. Princeton, N.J.: Princeton University Press.
- Fitzpatrick, Sheila. 1982. *The Russian Revolution, 1917-1932*. Oxford: Oxford University Press.
- Floor, Willem. 1983. "Change and Development in the Judicial System of Qajar Iran." In *Qajar Iran: Political, Social and Cultural Change, 1800-1925*, edited by E. Bosworth and C. Hillenbrand. Edinburgh: University of Edinburgh Press.
- Furet, François. 1981. *Interpreting the French Revolution*. Cambridge: Cambridge University Press.
- Goldstone, Jack A. 1982. "The Comparative and Historical Study of Revolutions." *Annual Review of Sociology* 8:187-207.
- . 1991a. "Ideology, Cultural Frameworks, and the Process of Revolution." *Theory and Society* 20, no. 4 (August): 405-53.
- . 1991b. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, N.J.: Princeton University Press.
- Halid Bey, Halil. 1909. "The Origin of the Revolt in Turkey." *Nineteenth Century and After* 65:755-60.
- Harcave, Sidney. 1964. *First Blood: The Russian Revolution of 1905*. New York: Macmillan.

- Healy, Ann Erikson. 1976. *The Russian Autocracy in Crisis, 1905–1907*. Hamden, Conn.: Archon Books.
- Hermasi, Elbasi. 1976. "Toward a Comparative Study of Revolutions." *Comparative Studies in Society and History* 18, no. 2 (April): 211–25.
- Heyd, Uriel. 1961. "The Ottoman Ulema and Westernization in the Time of Selim III and Mahmud II." *Scripta Hierosolymitana* 96:63–96.
- Hidayat, Mahdi Quli. 1982. *Khatirat va Khatarat*. Tehran: Kitabfurushi-i Zuvvar.
- Hosking, Geoffrey A. 1973. *The Russian Constitutional Experiment: Government and Duma, 1907–1914*. Cambridge: Cambridge University Press.
- Hosking, Geoffrey A., and Roberta Thompson Manning. 1979. "What Was the United Nobility?" In *The Politics of Rural Russia, 1905–1914*, edited by Leopold H. Haimson. Bloomington: Indiana University Press.
- Kasravi, Ahmad. 1951. *Tarikh-i Mashrutah-i Iran*. Gilan: Amir Kabir.
- Keep, John. 1968. "1917: The Tyranny of Paris over Petrograd." *Soviet Studies* 20, no. 1 (July): 22–35.
- Kimmel, Michael S. 1990. *Revolution: A Sociological Interpretation*. Philadelphia: Temple University Press.
- Kirman, Ahmad Majid al-Islam. 1972. *Tarikh-i Inhilal-i Majlis: Fasli az Tarikh-i Inqilab-i Mashrutiyyat-i Iran*. Isfahan: Intisharat-i Danishgah-i Isfahan.
- Kirman, Muhammad Nazim al-Islam. 1983. *Tarikh-i Bidari-i Iranian*. 2 vols. Edited by 'Ali Akbar Saidi Sirjani. Tehran: Intisharat-i Agah, Intisharat-i Nuvin.
- Kissling, Hans Joachim. 1954. "The Sociological and Educational Role of the Dervish Orders in the Ottoman Empire," edited by G. E. von Grunebaum. *American Anthropologist*. Memoir 76.
- Knight, E. F. 1909. *The Awakening of Turkey: A History of the Turkish Revolution*. London: John Milne.
- Kochan, Lionel. 1966. *Russia in Revolution, 1890–1918*. London: Weidenfeld & Nicholson.
- Kuran, Ahmet Bedevi. 1945. *İnkılap Tarihimiz ve "Jön Türkler"*. Istanbul: Tan Matbaası.
- . 1948. *İnkılap Tarihimiz ve İttihad ve Terakki*. Istanbul: Tan Matbaası.
- Lambton, Ann K. 1953. *Peasants and Landlords in Persia*. London: Oxford University Press.
- . 1963. "Persian Political Societies, 1906–1911." *St. Antony's Papers*, no. 16. Middle Eastern Affairs, no. 3.
- Lefebvre, Georges. 1947. *The Coming of the French Revolution*, translated by R. R. Palmer. Princeton, N.J.: Princeton University Press.
- Levin, Alfred. (1940) 1966. *The Second Duma: A Study of the Social-Democratic Party and the Russian Constitutional Experiment*, 2d ed. Hamden, Conn.: Archon Books.
- Lewis, Bernard. 1953. "The Impact of the French Revolution in Turkey." *Journal of World History* 1, no. 1 (July): 105–25.
- . 1961. *The Emergence of Modern Turkey*, 2d ed. Oxford: Oxford University Press.
- Manning, Roberta Thompson. 1979. "Zemstvo and Revolution: The Onset of the Gentry Reaction, 1905–1907." In *The Politics of Rural Russia, 1905–1914*, edited by Leopold H. Haimson. Bloomington: Indiana University Press.
- . 1982. *The Crisis of the Old Order in Russia: Gentry and Government*. Princeton, N.J.: Princeton University Press.
- Mardin, Şerif. 1962. *The Genesis of Young Ottoman Thought: A Study in the Modernization of Turkish Political Ideas*. Princeton, N.J.: Princeton University Press.
- McCarthy, John D., and Meyer N. Zald. 1977. "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82:1212–41.

- Nadi, Yunus. 1327/1909. *İhtilal ve İnkılab-i Osmani*. Istanbul: Cihan.
- Nashat, Guity. 1982. *The Origins of Modern Reform in Iran, 1870–1880*. Urbana: University of Illinois Press.
- Pettee, George S. 1938. *The Process of Revolution*. New York: Harper & Brothers.
- Pintner, Walter McKenzie, and Don Karl Rowney, eds. 1980. *Russian Officialdom: The Bureaucratization of Russian Society from the Seventeenth to the Twentieth Century*. Chapel Hill: University of North Carolina Press.
- Raeff, Marc. 1984. *Understanding Imperial Russia: State and Society in the Old Regime*. New York: Columbia University Press.
- Ragin, Charles C. 1987. *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California Press.
- Ra'iss Tousi, Reza. 1988. "The Persian Army, 1880–1907." *Middle Eastern Studies* 24, no. 2 (April): 206–29.
- Russel, D. E. H. 1974. *Rebellion, Revolution, and Armed Forces: A Comparative Study of Fifteen Countries with Special Emphasis on Cuba and South Africa*. New York: Academic Press.
- Sabahaddin, M. 1908a. *Teşebbüs-i Şahsi ve Tevsi'-i Me'zuniyet Hakkında Bir İzah*. Istanbul: Cihan.
- . 1908b. *Teşebbüs-i Şahsi ve Tevsi'-i Me'zuniyet Hakkında İkinci Bir İzah*. Istanbul: Mahmud Bey Matbaası.
- Sewell, William H., Jr. 1985. "Ideologies and Social Revolutions: Reflections on the French Case." *Journal of Modern History* 57, no. 1 (March): 56–85.
- . In press. "Three Temporalities: Toward an Eventful Sociology." In *Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Shaw, Stanford J., and Ezel Kural Shaw. 1977. *History of The Ottoman Empire and Modern Turkey*. Vol. 2: *Reform, Revolution, and Republic: The Rise of Modern Turkey, 1808–1975*. Cambridge: Cambridge University Press.
- Shlapentokh, Dimitri. 1988. *The French Revolution in Russian Intellectual and Political Life, 1789–1922*. Ph.D. dissertation. University of Chicago, Department of History.
- Skocpol, Theda. 1976. "France, Russia and China: A Structural Analysis of Social Revolutions." *Comparative Studies in Society and History* 18, no. 2 (April): 176–210.
- . 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
- . 1985. "Cultural Idioms and Political Ideologies in the Revolutionary Reconstruction of State Power: A Rejoinder to Sewell." *Journal of Modern History* 57, no. 1 (March): 86–96.
- Skocpol, Theda, and Margaret Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22:174–97.
- Stone, Lawrence. 1966. "Theories of Revolution" *World Politics* 18, no. 2 (January): 159–76.
- Taqizadah, Sayyid Hasan. 1957. *Zaminah-i Inqilab-i Mashrutiyyat-i Iran: Sih Khatabah*. Tehran: Intisharat-i Gam.
- Tilly, Charles. 1978. *From Mobilisation to Revolution*. New York: Random House.
- Tilly, Charles, Louise Tilly, and Richard Tilly. 1975. *The Rebellious Century, 1830–1930*. Cambridge, Mass.: Harvard University Press.
- Tunaya, Tanık Zafer. 1952. *Türkiyede Siyasi Partiler, 1859–1952*. Istanbul: Doğan Kardeş Yayınları.
- . 1984. *Türkiyede Siyasal Partiler, Cilt I: İkinci Meşrutiyet Donemi, 1908–1918*. Istanbul: Hurriyet Vakfı Yayınları.
- Turkaman, Muhammad. 1983. *Rasa'il, I'lamiyah-ha, Maktubat, . . . , va Ruznamah-i Shaykh Fazel Allah Nuri*, vol. 1. Tehran: Mahtab.

- Verner, Andrew M. 1990. *The Crisis of Russian Autocracy: Nicholas II and the 1906 Revolution*. Princeton, N.J.: Princeton University Press.
- Weber, Max. 1946. "The Social Psychology of the World Religions." Pp. 267–301 in *From Max Weber: Essays in Sociology*, translated and edited by H. H. Gerth and C. Wright Mills. New York: Oxford University Press.
- Yaney, George L. 1973. *The Systematization of Russian Government: Social Evolution in the Domestic Administration of Imperial Russia, 1711–1905*. Urbana: University of Illinois Press.

## OTHER SOURCES

### Newspapers and Official Publications

The dates that appear in conjunction with the title and place of publication correspond to the time periods used for this article and not necessarily to the entire duration of publication.

### *The Ottoman Empire and Turkish Republic*

- Düstur* (in text as *Düstur2*). Second ser. (tertib-i sani), vol. 1, 1908–9. Istanbul: Matbaa-i Osmaniye, 1911. (This is the official collection of laws, regulations, and decrees, 1908–9).
- Fikir Hareketleri*. Periodical, Istanbul, 1935.
- İkdam*. Semiofficial newspaper of the Liberals, Istanbul, 1908–9.
- Şura-yi Ümmet*. Young Turk newspaper, Paris and Cairo, 1902.
- Tanin*. Semiofficial newspaper of the CUP, Istanbul, 1908–9.
- British Documents on Foreign Affairs: Reports and Papers from the Foreign Office Confidential Print*. Part 1: *From the Mid-Nineteenth Century to the First World War*. Ser. B: *The Near and Middle East, 1856–1914*. Vol. 20: *The Ottoman Empire under the Young Turks, 1908–1914*, edited by David Gillard. Frederick, Md.: University Publications of America, 1985.

### *Iran*

- Anjuman*. Newspaper of Tabriz Provincial Committee, Tabriz, 1907.
- Habl al-Matin*. Constitutionalist newspaper, Tehran, 1907–8.
- Majmu'ah-i Musavvat-i Advar-i Avval va Duvvum-i Qanun-Guzari-i Majlis-i Shawra-yi Milli* (in text as *Musavvat I-II*). Tehran: Idarah-'i Qavanin va Matbu'at-i Majlis-i Shawra-yi Milli. (The official collection of laws and regulations for the first and second legislative periods, 1906–11).
- Musavat*. Constitutionalist newspaper, Tehran and Tabriz, 1907–9.
- Ruznamah-i Majlis* (in text as *Majlis1*). Record of parliamentary debates during the first legislative period, Tehran, 1906–8.
- Sur-i Israfil*. Constitutionalist newspaper, Tehran, 1907–8.

# Leaps of Faith: Shifting Currents in Critical Sociology of Education<sup>1</sup>

Scott Davies

*McMaster University*

The sociology of education is fraught with noncommunicating paradigms, particularly between “critical” and “mainstream” approaches. This article outlines and explains this paradigm split and the resulting shifts of critical theories from “reproduction” to “resistance” to “post-Marxist” versions, then explains how these approaches differ from the mainstream and what their internal dynamic is. The article argues that the “immanentism”—the practice of asserting a necessary movement of history that confers subordinate groups with objective interests in radical change—of critical approaches is the source of theoretical difference and that the trajectory of critical theories reflects shifting conceptions of how this change will occur.

## INTRODUCTION: PARADIGMS AND THE SOCIOLOGY OF EDUCATION

Sociology is riddled with paradigms. Various subareas of the discipline are fragmented into self-enclosed, noncommunicating camps, perhaps most egregiously between “critical” and “mainstream” sociologists.<sup>2</sup> There are conflicting assessments of this state of affairs. Randall Collins

<sup>1</sup> I would like to thank Julian Tanner, Bonnie Erickson, Neil Guppy, Axel van den Berg, Rick Ogmundson, Dave Tindall, Luc Theriault, and the *AJS* reviewers for their valuable comments. Funding for this research came from a Social Sciences and Humanities Research Council of Canada doctoral fellowship and a SSHRC postdoctoral fellowship. Please direct all correspondence to Scott Davies, Department of Sociology, McMaster University, Hamilton, Ontario, Canada L8S 4M4. E-mail: [daviesrs@mcmaster.ca](mailto:daviesrs@mcmaster.ca)

<sup>2</sup> By “critical” I refer to works that trace a direct lineage to Marxism. My use of the term is not to be confused with the particular tradition associated with the Frankfurt school, though many of these sociologists are influenced by that tradition. Nor does it include sociologists such as Pierre Bourdieu or Basil Bernstein, who, though somewhat influenced by Marxism, stray far from its central core. I do not use the term “Marxist,” because, as I describe below, one of the key variants is avowedly “post-Marxist.” By “mainstream” I refer to the majority of sociologists who do not claim to be part of the critical tradition. This negative definition is used because it represents how critical theorists mark territory within the field.



(1986, 1989) laments this fragmentation, for he sees it as hindering sociology's potential for cumulative knowledge. To the contrary, "post-positivists" (e.g., Alexander 1982) encourage us to accept this situation and, in the prevailing mood of the philosophy of science, recognize the incompatible worldviews that underlie these divisions.

The sociology of education is no exception. While class inequality in schooling is the most enduring topic in this subdiscipline, many researchers question whether this area has enjoyed continuous advances. In 1977, for instance, acclaimed reviewers (Karabel and Halsey 1977; Rist [1977] 1990) highlighted the limitations posed by the prevailing rivalries between competing traditions. But they remained optimistic and urged researchers to transcend these divisions by synthesizing promising micro and macro approaches. Not surprisingly, those calls have been only partially heeded. On the one hand, earnest efforts to bridge structure and agency have emerged since the mid-1970s, most prominently in critical accounts of class inequality in education. Paul Willis's (1977) massively influential *Learning to Labour* has spawned an imposing corpus of U.S., British, Canadian, and Australian ethnographies and theoretical commentaries on school resistance.<sup>3</sup> But paradoxically, this literature exemplifies the deepening of paradigm divisions. Willis and his followers engaged in little explicit contact with earlier non-Marxist subcultural theories, despite the common empirical terrain and striking similarities between these traditions, as noted by various commentators (e.g., Cohen 1980; Frith 1985; Tanner 1988). Indeed, critical theorists rarely cite any preexisting or contemporary mainstream sociology of education. The few citations one uncovers are brief and dismissive of this "old" or "traditional" sociology, implying that critical work represents a bold advance (e.g., Connell et al. 1982, pp. 24–28; Wexler 1987, pp. 27–32; Holland and Eisenhart 1990, pp. 26–27; Aronowitz and Giroux 1991, p. 13; McLaren 1989).

Critical theorists justify these paradigm divisions by proclaiming their project to be inherently irreconcilable with the concerns of positivists. This incompatibility is said to be rooted in, for some, the distinctiveness of the ethnographic method (Willis 1983), for others, the guiding orientation of praxis (Giroux 1983), and for others still, the use of an interpretivist framework (Wexler 1987). Each of these is presumed to generate

<sup>3</sup> Popular examples that either invoke or discuss Willis's framework include Anyon (1981), McRobbie (1981), Connell et al. (1982), Giroux (1983), Aggleton and Whitty (1985), Apple (1985), Aronowitz and Giroux (1985, 1991, 1993), Gaskell (1985), McLaren (1986, 1989), Macleod (1987), Wexler (1987), Liston (1988), Eckert (1989), Holland and Eisenhart (1990), Tanner (1990), Weis (1990), Leong (1992), and Mehan (1992).

vastly different knowledge than that found in the mainstream. As a result, the debates engaged by critical researchers remain largely within their own paradigm. Yet this insularity has itself promoted a remarkable trajectory of approaches that bear the common stamp of critical theory. This trajectory has no counterpart in the mainstream, for it is distinguished by some extraordinary about-faces in substantive claims and interpretations. Their depictions of quintessential young working-class males provide a particularly revealing example.

The initial critical approach was reproduction theory. By proposing that schools were pivotal for the functioning of capitalism and that class disparities resulted from the system's needs for stratified labor, theorists such as Bowles and Gintis (1976) portrayed working-class youth as passive victims of schools' sorting mechanisms and manipulative socialization.

Though this approach was heralded by its supporters as a "great advance" that put "liberal-pluralist accounts of education clearly on the run" (Aronowitz 1981), it was replaced in relatively short order by the generation of resistance theories. Willis and his followers insisted that class inequalities were not unproblematically reproduced by schools but were contested by working-class youth. Resistance theorists inverted the reproductionist image of these youth by championing school rejectors as harbingers of an oppositional culture with ominous political potential.

More recently, a post-Marxist variant is supplanting resistance theory as the predominate critical approach to education (e.g., Aronowitz and Giroux 1985, 1991, 1993; Wexler 1987; Giroux and McLaren 1989). Post-Marxists claim to eschew some key Marxian tenets, and extend their scope beyond social class to include race and gender. In doing so, they have altered the depiction of working-class males, who are now increasingly distrusted as potential adherents of the New Right, even though their actions have scarcely changed since the late 1970s (e.g., Apple 1989; Weis 1990).

Though reproductionists, resistance theorists, and post-Marxists share a common lineage, on the surface they are diametrically opposed, judging by their starkly contrasting depictions of young working-class males. Yet critical theorists do not perceive any problems here and actually regard this trajectory as a culmination of increasing sophistication (e.g., Aronowitz and Giroux 1991; McLaren 1989). Consequently, critical approaches have proliferated while becoming increasingly dismissive of the mainstream.

#### WHAT UNDERLIES THE TRAJECTORIES OF THESE THEORIES

These theoretical twists and turns—the entrenched divisions between critical and mainstream paradigms and the remarkable development

*within* the critical approach—are the objects of interest in this article. Rather than attempting an exhaustive review of the sociology of education, I focus on competing explanations of class inequalities in education that have been offered by critical and mainstream researchers. In particular, I follow two primary lines of inquiry.

First, I investigate what theoretical principles mark the boundary between critical and mainstream approaches and what underlying rules continue to set critical approaches apart from the mainstream while allowing them to remain within their paradigm when they change their form. By comparing approaches that share a common explanatory object, I attempt to uncover which differences reflect true paradigm boundaries, in terms of incompatible assumptions, presuppositions, modes of reasoning, and the like, as opposed to less profound differences in emphasis or focus. These are useful inquiries. Even if one accepts the postpositivist creed that paradigms cannot be judged by a universal criterion, it is valuable to uncover just how they differ in terms of the particular grounds on which their respective adherents disagree. Rather than unquestioningly accept claims of incommensurability, such an exercise may reveal more compatibility than previously thought (van den Berg 1993). Discovering compatibility may promote greater dialogue and perhaps help revitalize research through greater openness and avoidance of repetition and theoretical parochialism, all of which could expand opportunities for synthesis and the production of cumulative knowledge.

Second, I investigate what underlies the major transformations within critical theory—from reproduction to resistance to the now-ascending post-Marxist—position and what internal pressures induce these changes. Similar theoretical divisions and trajectories have been analyzed in other areas of sociology, such as studies of the state (van den Berg 1988) and industrial sociology (Tanner, Davies, and O'Grady 1992), so it is worth examining the continuities or discontinuities between these areas and the sociology of education, especially since the latter, unlike the former, lacked a rich Marxist tradition before the late 1960s. Studying trends in this subdiscipline highlights the development of paradigms in latter-day sociology. Further, unlike those explorations in theories of the state and labor process, I include the more recent development of post-Marxism.

I begin by outlining the major tenets of older mainstream approaches—cultural deprivation theory, subcultural theory, and status attainment research—to provide a useful reference point. I then outline more fully the shifts among critical theories. Having established this background, I compare the resistance and subcultural literatures to explore what, if anything, distinguishes these approaches. Following that, I attempt to explain the trajectory of critical approaches. I contend that the critical approaches' immanentism—the practice of asserting a neces-

sary movement of history that confers subordinate groups with objective interests in radical change—is the actual source of their theoretical difference and that the trajectory of critical theories reflects shifting conceptions of how this change will occur and who will bring it about. These arguments confront critical theorists' own claims of what distinguishes them from the mainstream.

#### MAINSTREAM RESEARCH

An early effort to explain educational inequalities was the widely disputed cultural deprivation theory of the 1950s and 1960s. Grounded in functionalist modernization theory, this approach deemed various value orientations to be necessary in an industrial society and traced educational inequalities to the failure of lower socioeconomic groups to embrace those values. The purported lack of "middle-class" traits—such as individualism, need for achievement, competitiveness, ability to defer gratification, and future orientation—among lower socioeconomic groups was considered the primary source of educational disparities (see Banks 1976; Hyman 1953).

Complementary to this approach was subcultural theory. Albert Cohen (1955) explained lower-class males' propensity for delinquency by borrowing assumptions from the deprivationists. He argued that, though these boys desired the same material goals as their middle-class counterparts, their cultural deficits limited their educational chances. Their subsequent status deprivation motivated them to establish an antischool subculture as an alternative source of status. This theory was further developed in the United States and Britain. Stinchcombe (1964) qualified the importance of class background by emphasizing how a lack of school success, regardless of one's socioeconomic background, prompted rebellion by making school appear irrelevant for one's future. He also stressed that expressions of rebellion differed by gender. While male emblems of adulthood comprised smoking, drinking, driving, and sexual bravado, unsuccessful girls were seen to adopt adult symbols of femininity that were less at odds with school codes of conduct, such as dating and wearing engagement rings. Consequently, Stinchcombe claimed that exhibitions of rebellion followed gender lines more than class divisions. Later in Britain, Hargreaves (1967), Lacey (1970), Downes (1966), and Murdock and Phelps (1973) similarly drew on Cohen to explain antischool subcultures but modified them for the British context. Hargreaves and Lacey outlined how streaming systems prompted a mounting polarization between pro- and antischool subcultures. Downes and Murdock and Phelps argued that working-class youths' low prior expectations and early streaming did not promote extreme reactions against school but

rather fostered perceptions of school as irrelevant. Thus, delinquency reflected boredom more than a protest against thwarted ambitions.

Subcultural theorists presented these substantive empirical claims. All agreed with Cohen's claim that, if educational opportunities are blocked, students experience status deprivation. These youth thus dissociate from school culture by forming subcultures, an antagonistic machismo among males and an orientation toward domesticity among females. These theorists disagreed largely over whether this dissociation is better characterized as youthful reactions to blocked opportunity and status deprivation or as an assertion of a working-class culture that is imported into the school.

Many sociologists in the 1970s rejected cultural deprivation approaches as part of a general renunciation of structural functionalism. Critics accused the deprivationists of ignoring the arbitrary middle-class biases within schools that disadvantaged working-class students and questioned the explanatory emphasis deprivationists placed on cultural over structural factors (e.g., Valentine 1968; Ryan 1971). These antideprivationists contended that ostensible cultural differences among socioeconomic groups reflected differing adaptations to imposed opportunity structures, as opposed to any inherent cultural dissimilarity. For instance, Ryan (1971, pp. 124–25) claimed that every second-generation American, regardless of class or race, participates in a common culture. To claim that lower socioeconomic groups have a separate culture, he argued, is a "manifest absurdity," for there was no evidence that the poor perceived their way of life as preferable to the dominant culture. Moreover, Ryan dismissed cultural deprivation traits as "minor phenomena," most of which he felt were shared by middle-class youth. But perhaps most important, both Ryan and Valentine contended that to place explanatory priority on alleged class cultural differences "blamed the victim," for it deflected blame from the real sources of the problems in schools, such as tracking and biased teacher expectations.

Meanwhile, subcultural theory was similarly criticized for overstating cultural differences among different classes of youth. Though not as ideologically charged as the protests against the deprivationists, critics complained that subcultural theory overpredicted working-class school failure and delinquency. These commentators argued that delinquency was better seen as an episodic "moral holiday" that drew on "subterranean" values shared widely among most youth (see Frith 1985; Brake 1985; Downes and Rock 1988).

Since the mid-1970s, the predominate mainstream approach in American sociology of education has been status attainment research. Stemming from the pioneering work of Blau and Duncan (1967) and other mobility research, this literature has been further refined in the Wisconsin

models (e.g., Sewell and Hauser 1980). It has rigorously employed survey data to trace the process of individual educational attainment. A particular focus has been on assigning relative weights for ascriptive versus achieved determinants of school success and assessing school's role in promoting meritocracy and social mobility. Since the late 1960s and early 1970s, this research tradition has continued to document socioeconomic disparities in school education. Status attainment research remains ascendant today in American sociology of education but is increasingly broad and diverse in its substantive concerns and styles of modeling.

#### THE "NEW SOCIOLOGY OF EDUCATION": FROM REPRODUCTION TO RESISTANCE TO POST-MARXISM

The criticisms of these mainstream approaches were implicit in the rupture in the sociology of education of the 1970s. Reproduction theory, the initial Marxist approach to education, incorporated the view of middle-class-biased schools but took it one step further. Bowles and Gintis (1976) and Apple (1979) implicated schools in the maintenance and legitimation of capitalism. For them, schools' veil of teaching neutral skills and enabling fair competition hid the reproduction of labor for relatively fixed class structures. Hierarchical streaming was seen to stratify and divide the labor force and to promote a capitalist "hidden curriculum" that repressed the potential for the development of a mass prosocialist consciousness.

Reproductionists were influenced by the prevailing mainstream research in varying ways. Bowles and Gintis borrowed the antideprivationists' resolute emphasis on opportunity structures to argue that schools promoted passive adaptations among working-class students. It is interesting that Bowles and Gintis offered few criticisms of either subcultural theory or the cultural deprivationists. Rather, they outlined the structural determinants that "cooled out" students into accepting predetermined subordinate roles. The reproductionists did, however, give more attention to status attainment and mobility research. They shared the concern with social mobility and readily borrowed this research to draw attention to class inequalities in education, while adding some original statistical evidence. However, what was not to Bowles and Gintis's liking was the mainstreamers' belief that these inequalities could be remedied with further reform, at least in principle (see Halsey, Heath, and Ridge 1980; Coleman and Husen 1985). In contrast, Bowles and Gintis reinterpreted the same data as proving that inequalities were endemic to capitalist schools. They attacked "the conventional wisdom . . . that within the 'free enterprise' system of the United States, significant social progress can be achieved" (pp. 5-6) and concluded that "the education system

is a major arena in which the democratic and productive potential of capitalism confronts its limits" (p. 14).

Bowles and Gintis thus labeled the mainstreamers "liberal" (an unflattering epithet in the hands of radicals) because of their optimism toward the possibility of eventually reducing disparities, though they freely admitted that this posed a difficult task. Bowles and Gintis's criticisms were not based on unique empirical data or procedures. Rather, the reproductionist position was distinguished by the conviction that educational reforms were rendered futile by the surrounding capitalist environment and that a radical alternative was thus required. Consequently, that was the core of this paradigm break.<sup>4</sup> But what was never really answered was why the reproductionists regarded reforms to be hopeless and put such stock in vague revolutionary alternatives.

This unabridged version of reproduction theory was relatively short-lived. Though many Marxists remained sympathetic to the tenet that education essentially served capitalist functions, reproductionism was criticized for being too mechanical, thereby failing to offer a suitable account of the processes that led to unequal outcomes (Willis 1977; Giroux 1983; Apple 1985; Connell et al. 1982). The antidote was sought in a renewed emphasis on "agency," "opposition," and "contestment," to revise the portrayal of untrammelled capitalist power and working-class passivity. To inject dynamic elements of class struggle into a neo-Marxist theory of education, Marx's dictum "men make history but not as they please, in circumstances not of their choosing" became an essential credo and the intellectual plank for justifying a renewed focus on agency while avoiding an unsavory voluntarism (e.g., Giroux 1983, p. 259).

Reproduction theory gradually evolved into resistance theory. Relations between school and working-class youth were now characterized not as passive but as "guerilla warfare" (CCCS 1981). Resistance theorists regarded school opposition as a youthful expression of a proletarian culture. While acknowledging varieties of working-class student groups, such as Willis's "lads and ear'oles," and the importance of race and

<sup>4</sup> Critical theorists have also reprimanded status attainment research for being overly descriptive and narrow. Whether or not this criticism is valid, it fails to mark paradigm boundaries, since similar criticisms have been repeatedly voiced within the mainstream (e.g., Campbell 1983), and since status attainment research has grown increasingly diverse and now tackles a wider variety of issues. Moreover, since the heyday of reproductionism, most radicals have ignored status attainment research or dismissed it altogether (e.g., Wexler 1987). In fact, one could argue that this situation has created a type of cultural lag. Unlike Bowles and Gintis, most contemporary critical theorists appear to be largely unfamiliar with mainstream research, such that their characterizations of the mainstream seem to pertain more to the 1970s than to current writings.

gender as pivotal sources of variation, these researchers nonetheless asserted that school resistance represents working-class culture in no uncertain terms.

While it may appear that this insistence on incessant class struggle in education is not easily reconciled with the realities of continual class inequalities and lack of revolutionary fervor—the very outcomes that prompted the determinism of reproduction theory—the resolution is found in the notion of “reproduction through resistance.” Simply put, working-class students are said to condemn themselves to working-class futures by affirming their autonomous culture. Their “partial insights” into school’s middle-class and capitalist biases motivate a rebellion that dooms them to educational failure. For boys, this revolt is characterized by confrontations with teachers, petty delinquency, and truancy. The feminized version of this thesis (see McRobbie 1981; Apple 1985) held that working-class females’ opposition to school is less aggressive, but these researchers discovered “resistance” in the girls’ embrace of an exaggerated femininity and infatuation with romance. This resistance was seen to symbolically taunt the “middle-class” ideal of a diligent female student but also to entrap these girls into the roles of marriage and motherhood. Thus, for working-class males and females, the essential irony is that in contesting their subordination they reproduce themselves as manual laborers and homemakers (Willis 1977; Giroux 1983; Apple 1985).

A new critical theory of education has recently emerged that does not center predominately on class nor draw its inspiration solely from Marxism. Post-Marxism also draws on developments in feminism, literary theory, postmodernism, and poststructuralism (McLaren 1989; Aronowitz and Giroux 1985, 1991, 1993; Wexler 1987; Giroux and McLaren 1989; Weis 1990).<sup>5</sup> Probably the first major post-Marxist work on education was Aronowitz and Giroux’s *Education under Siege* (1985). With sections and chapters entitled “Reconstructing Marxist Discourse: Beyond Class and Economism” and “Radical Pedagogy and the Legacy of Marxist Discourse,” this book claimed to confront the limits of the prevailing Marxist sociology of education and hoped to rebuild radical theory. Three major claims distinguish post-Marxists from their predecessors.

First, they claim to avoid Marxism’s exclusive focus on social class. Their new starting point is to view class, gender, and race as irreducible categories of subjectivity and social relations. By not privileging class

<sup>5</sup> I do not focus on “critical pedagogy” per se, which is a philosophy of teaching that is connected to post-Marxism. Rather, I outline the implicit explanations used by post-Marxists to comprehend how inequalities are generated.



over these other social categories, they proclaim a more inclusive theory of domination, resistance, and educational equality. Second, post-Marxists look to new groups for leftist political potential. While Marxist sociologists of education implicitly designated working-class school rejectors as potential allies of a militant labor movement, post-Marxists look to the new social movements as more promising agents of radical transformation.<sup>6</sup> They urge educators to link school resistance to feminist, environmental, antiracist, and gay liberation struggles (e.g., Aronowitz and Giroux 1985; Wexler 1987; Giroux 1989; Weis 1990). Third, post-Marxists reject "positivistic" premises that they claim are latent even within Marxism. Instead they embrace interpretivism,<sup>7</sup> which they claim is the only means to uncover the socially constructed nature of knowledge (Wexler 1987; Aronowitz and Giroux 1985).

To summarize, since its rupture from the mainstream, critical approaches have evolved from reproduction theory to resistance theory to post-Marxism. But in what concrete ways do these theories differ from the mainstream? What serves as the paradigmatic boundary? To answer these questions, I compare two approaches—resistance theory and subcultural theory—that cover a strikingly similar empirical territory. In my comparison, I use the substantive claims of the subculturalists and of Willis in particular, whose work critical theorists still regard as the quintessential depiction of working-class relations with schooling (e.g., McLaren 1989; Holland and Eisenhart 1990; Aronowitz and Giroux 1993).

#### HOW RESISTANCE THEORY DIFFERS FROM MAINSTREAM APPROACHES

Do these theories differ markedly in their substantive claims? Consider the actual empirical content of Willis's (1977) celebrated book. He describes the crystallization of an informal group of the "lads" in the lower streams of the school during the year before they are to leave for work. Their distinct subculture is said to be expressed in their precocious sexuality, drinking, smoking, disobedience, fighting, avoidance of school work, and dislike of teachers. Moreover, they believe they possess a worldly

<sup>6</sup> Indeed, Paul Willis, a pioneering chronicler of working-class youth resistance, has entered a post-Marxist phase in his *Common Culture* (1990). He now acknowledges the commonalities of British youth culture across class boundaries and sees radical potential in youths' creative consumerism.

<sup>7</sup> Interpretivists question scientific ways of thinking. Rather than attempting to describe and explain an objective reality, they instead focus on uncovering ways in which reality is represented and interpreted by different groups of people in various contexts.

superiority over teachers and school conformists as they celebrate their adolescent masculinity, a revelry that unfortunately also contains racist, sexist, and anti-intellectual elements.

Compare these descriptions to subcultural classics such as Stinchcombe (1964) and Hargreaves (1967). These researchers similarly documented the emergence of tightly knit male groups that subverted authority with an exaggerated masculinity symbolized by petty delinquency, fighting, and sexual bravado. Hargreaves, like Willis, discovered these acts to be most predominate in low streams near the end of their compulsory schooling. Both Hargreaves and Stinchcombe viewed these phenomena as reactions to streaming and school failure, which serve to weaken youths' stakes in conformity and impose an inglorious adolescent status. Furthermore, Stinchcombe, foreshadowing McRobbie (1981), chronicled how failing females embraced symbols of romance in their attempt to claim adult status.

In what ways do resistance and subcultural theory thus differ? They identify almost identical manifestations of school opposition: emerging subcultures, petty delinquency, antagonisms toward teachers, perceptions of schooling as irrelevant and boring, and gender traditionalism. Though resistance theorists clearly attribute school opposition to be the province of working-class youth, and subcultural theorists disagree among themselves on this matter, this difference hardly represents a paradigmatic boundary, since class rates of participation in school rebellion can be, in principle, empirically tested (as done by Davies [1993, 1994a, 1994b]). However, resistance theory does differ by its markedly dissimilar interpretations of the same manifestations of school opposition. Whereas subcultural theorists portrayed school rebellion as a youthful expression of disappointment over thwarted goals or a fatalistic disengagement from school, the term "resistance" conveys something more vital and radical than the equivalent phrases in subcultural theory (Cohen 1980).

Indeed, Willis (1977) interprets these familiar expressions of student discontent as a form of class politics. For him, the lads' sentiments have a powerful affinity with shop-floor culture and as such are a youthful expression of the wider class culture. But more important, by equating counterschool culture with working-class culture, he presumes that it contains embryonic "insights" into various capitalist realities, such as credentialism, the falseness of the mobility ethic, and the meaninglessness of deskilled work. He reasons that since schools' bogus promises of equal opportunity help incorporate the working class into capitalism, the rejection of schooling subverts this illusory allure and signals a tacit ideological preference for collectivism over individualism. At the root of these unique

interpretations is the conviction that school rebellion embodies an albeit buried, partial, and poorly articulated critique of capitalism. In contrast to the subcultural theorists, who did not associate school rebels with any particular political orientation, resistance theorists detect a potential for a working-class movement with leftist undertones (Willis 1977; Corrigan 1979; Giroux 1983).

Resistance theorists thus transform many of the same student traits identified by the cultural deprivationists into favorable attributes. Whereas the deprivationists traced these traits to the strain of unequal opportunity, resistance theorists link them to the exigencies of capitalist production and see them as expressions of proletarian culture. For instance, working-class students' purported rejection of competitive individualism is said to signify impending class solidarity and a rejection of capitalist values. These youths' "anti-intellectualism" becomes a "rejection of the mental-manual division of labour" (Willis 1977) and "a celebration of their class-specific knowledge" (Apple 1985; see also McLaren 1989).

But on what grounds do these daring interpretations rest? Not surprisingly, they have prompted criticism. Resistance theorists have been accused of piling these dramatic interpretations over slender evidence, faulted for overlooking youths' conscious intentions, and criticized for condescendingly portraying nonoppositional youth as not "authentically" working class (Cohen 1980; Walker 1986; Hargreaves 1982; Frith 1985). In their defense, one popular claim among resistance writers is that these interpretations are the product of ethnography—their preferred method—which provides privileged insights (e.g., Connell et al. 1982; Willis 1983). But how convincing is it to argue that this method distinguishes the resistance literature, since the older sociology of education both in the United States (e.g., the work of Stinchcombe [1964]) and Britain (e.g., the works of Lacey [1970] and Hargreaves [1967]) possessed a substantial ethnographic literature? Acknowledging this, Henry Giroux (1983, pp. 98–99) writes:

Neo-Marxist accounts have not sacrificed theoretical depth for methodological refinement. That is, more recent Marxist studies have not followed the generally bland method of merely providing exhaustive descriptive analyses of the internal workings of the school. Instead, these perspectives—especially Willis (1977)—have attempted to analyze how determinant socioeconomic structures embedded in the dominant society work through the mediation of class and culture in shaping the lived antagonistic experiences of students at the level of everyday life. . . . It attempts to link social structures and human agency to explore the way they interact in a dialectical manner, representing a significant theoretical advance over functional-structuralist and interactionist accounts.

Yet, non-Marxist field researchers certainly share these desirable if abstract goals and claim to have enjoyed some success (e.g., Hammersley 1985; Hargreaves 1982; Woods 1983; Mehan 1992).

The question remains, What is the particular point of departure of resistance ethnography? Recognizing this quandary, Giroux (1983, p. 100) continues:

These critics belabor neo-Marxist studies for not drawing on or conducting empirical studies grounded in the problematic of liberal social theory. The point of course, is that in redefining the nature of educational theory, the struggle will not be over the use of data or types of studies conducted. The real battle will be over the theoretical frameworks in use. . . . In short, neo-Marxist studies on resistance have performed an important theoretical service by reinserting empirical work into the framework of critical theory.

Consequently, what spearheads this "great advance" is not empirical method per se but the use of critical theory to interpret classroom activities. This admission makes more sense, for to claim that critical empirical methods are inherently incompatible with positivist methods is difficult to reconcile with the existence of the entire analytical Marxism school, represented by luminaries such as Erik Olin Wright, Jon Elster, and Adam Przeworski, who test hypotheses and use survey methods (for a discussion of analytical Marxism in the sociology of education, see Liston [1988]). Moreover, avowed critical theorists have embraced Marxist survey research when its conclusions are deemed palatable (e.g., Harp's [1991] and Morrow's [1985] endorsements of Livingstone's [1983] work).

Giroux (1983) has also contended that what makes the resistance approach unique is its "praxis," or orientation to progressive political ends. But this contention ignores that much mainstream sociology of education is policy oriented and that many educational policies since the 1960s have been beneficial for underprivileged youth, as admitted by Aronowitz and Giroux (1985, 1991, 1993). To claim that praxis distinguishes these critical approaches, one would need to demonstrate that critical approaches are superior in their beneficial effects. In light of the paucity of research demonstrating the preeminence of "critical pedagogy" over mainstream progressive education, however, this declaration must rely on some unspoken theoretical assumption.

The distinctiveness of resistance theory and the critical paradigm as a whole thus remain elusive. Resistance theory offers few original substantive empirical claims and does not employ an original empirical method. This lack of empirical difference would seem to draw attention to its theoretical presuppositions, since it largely reinterprets familiar data in a neo-Marxist framework. The question that then arises is What is it

about resistance theory that generates such divergent interpretations and thus marks the boundary from mainstream sociology of education?

### IMMANENTISM

I argue that what continually divides mainstream and critical sociology of education is the latter's successive application of the "immanent core" of Marxism (van den Berg 1988, 1993).<sup>8</sup> As van den Berg notes, Marxism has never been merely a framework for analyzing capitalist society that expresses a preference for revolutionary change. Rather, the distinctiveness of Marxism is rooted in its philosophy of history. Marx and Engels viewed history as having a meaning: a Hegelian march of reason toward its fulfillment. They wished to demonstrate that socialism was not merely a desirable state of affairs but was guaranteed by forces inherent in the present society. Their entire oeuvre was oriented to show, through philosophical reflection and historical analysis, how the will of history was embodied in various societal struggles and contradictions. Thus, capitalism not only dominates people but also *necessarily* prepares the agent of its demise—hence the proletariat's mission to destroy capitalism and usher in socialism. This philosophy implicitly attributes "objective" interests to designated agents of change. Traditional Marxists portray the working class as inherently compelled to oppose capitalism and embrace an unspecified conception of "socialism" (Goldthorpe 1988; Marshall 1983; van den Berg 1988).

As such, social action is set within an underlying dramaturgy of polarized struggle between capital and a socialist-bound working class. The working class is viewed through the lens of class formation, which venerates the configuration of cultural attributes that Marx assigned to the proletariat as necessary for overthrowing capitalism. Acts of solidarity, recognition of interests distinct from capitalists, and reluctance toward capitalist norms are comprehended as profound challenges to capitalism and yearnings for socialism.

However, what if workers do not appear to be traveling along a militant road? For Marx and the socialists of his time, prescribing revolutionary interests to the working class did not generate many theoretical entanglements, for they believed, rightly or wrongly, that working-class insurrection was close at hand. But the subsequent nonrevolutionary track record of the Western working classes has left modern neo-Marxists with a primordial choice: either to modify these precepts or, without

<sup>8</sup> My use of this term emphasizes the a priori attribution of "objective" interests to designated agents that is retained regardless of empirical evidence. For a somewhat different emphasis, see van den Berg (1988, 1993).

questioning these assumptions, to analyze why predictions have not come true. Obviously, Marxists have chosen the latter.

In contemporary sociology, this choice creates the gulf between mainstream and critical reasoning. Whether under the banner of reproduction, resistance, or post-Marxism, empirical observation is not about gathering observations to produce provisional estimates about the likelihood of this or that occurrence. On the contrary, critical researchers are fundamentally ambivalent toward disconfirming empirical evidence. When expectations fail to come to fruition, they respond with counterexplanations that usually involve identifying various blockages to the spontaneous emergence of a working-class socialist movement. Indeed the analysis of "ideology" and "hegemony," in all their myriad forms, has been Western Marxism's and critical theory's grand enterprise. From Lenin's theory of the limits of "trade-union consciousness," to Lukàcs's "reification," to Gramsci's "cultural hegemony," to the Frankfurt school's speculations on the entertainment industry, and to Althusser's "ideological state apparatuses," there are innumerable justifications for why the revolution has not occurred. The main tactic is to portray unexpected events as epiphenomenal detours in the pull of history, so that the assumption that workers are irresistibly drawn to socialism remains unquestioned; rather, the analysis shifts to detailing how various forces keep the spontaneous emergence of radicalism at bay. The result is that theoretical verification is possible only with the advent of a radicalized working class; otherwise the gap between theoretical expectation and empirical observation is filled by projecting expectations into the distant future. As many have commented, what is striking about the history of Marxist and critical thought is the proliferation of imaginative responses to the lack of proletarian revolutionary fervor in ways that save the theory (e.g., Parkin 1979; van den Berg 1988).

As a result, critical theorists have developed a few implicit models of how radical change is to occur. In Marx's time, the underlying conception was essentially chiliastic: the working class would eventually rise up and make the revolution (Wallerstein 1986). This dramatic vision of sudden transformation implies that forces of radical change are ripening, and hence revolution must be looming near. The Second International, however, was confronted with a conception of gradual reformism that did not envisage revolution in the foreseeable future (see van den Berg 1988). Orthodox Marxist critics who wished to discredit reformism portrayed it in a negative light, arguing that reforms siphon off potential for "true" change in an otherwise primed revolutionary movement.

Paving a middle road between these positions is the popular vision of Antonio Gramsci. His emphasis on the patient building of a cultural

counterhegemony is seen as offering an alternative between waiting for a doubtful sudden proletarian insurrection and accepting a bland non-Marxist reformism (van den Berg 1988). Indeed, Gramsci represents the preferred imagery of immanent change in the cultural studies and cultural Marxist frameworks that inform resistance theory. Cultural Marxism, most prevalent in the late 1970s and early 1980s, represents a stage when Marxists had grown disenchanted with organizations such as trade unions and labor parties (Crouch 1982). These Marxists turned their sights toward the daily practices of workers for the genuine insurgent potential they felt had been stifled in formal political institutions. Cultural Marxists thus probed the earthy mores and sociable interactions of workers and their communities for buried signs of incipient radicalism.

Yet, critical theorists have historically vacillated from embracing the working class as the major agent of radical transformation. For instance, by the mid-1980s, many critical sociologists lost faith in the prospect of an impending Marxist-inspired working-class movement. This loss of faith in the working class, coupled with mounting pressures to incorporate gender and race into their theories, led critical theorists to envisage a type of social transformation that placed less emphasis on class struggle. Yet, this de-emphasis on the working class is no aberration within the history of critical thought. Rather it is quite consistent within the legacy of Marxism itself, which Gouldner (1985) characterizes as a long-standing search for a historical agent who will radically transform society. As Gouldner argues, Marxists can see revolutionary agents in almost any strata, anywhere, anytime. While the industrial working class has been Marxism's traditional candidate, the proletariat's reformist tendencies over this century have prompted Marxists to vacillate from embracing this class to searching for a substitute. This tendency has spawned such varieties as Third World Marxism, new working class theory, Andre Gorz's theory of the *lumpenproletariat*, and most recently, new social movement theory. The rejection of the working class in favor of the new social movements can thus be seen as the latest instance of what Gouldner dubs "shopping for an agent." Now the quaint image of impending working-class solidarity is replaced by a radical democratic vision of a vast coalition of feminist, environmentalist, racial minority, and gay/lesbian movements (see Laclau and Mouffe 1987).

This discussion has relevance for contemporary critical sociology of education, because this literature has inherited these tendencies: a teleological view of history, the practice of attributing objective interests to subordinate groups, imposing a class formation model on working-class culture, being ambivalent toward disconfirming empirical evidence, and wavering between different conceptions of radical change and different

agents of change. Below it will be demonstrated that these immanentist assumptions erect the border between the two paradigms and set in motion the remarkable trajectory within the critical literature.

#### THE TRAJECTORY OF CRITICAL THEORIES AS SHIFTING CONCEPTIONS OF IMMANENT CHANGE

The above discussion of reproduction and resistance theories left two questions unanswered: Why were reproduction theorists so disparaging of existing school reforms? And what prompted resistance theorists' controversial interpretations of pupil subcultures? The first question can be further investigated by examining the logic implicit in Bowles and Gintis's (1976) attacks on the mainstream. The second question necessitates an examination of the effect of importing a class formation model in classroom research.

Reproductionism offers an intransigent and forceful tool for condemning capitalism. Despite its gloomy tone, it is tacitly optimistic, for it portrays capitalism as rife with contradictions and thus in need of constant maintenance to keep from bursting asunder. By regarding educational inequalities as necessary outcomes of the capitalist system, Bowles and Gintis (p. ix) assumed that "only in a . . . revolutionary framework can the schools fulfil what we take to be their . . . goal." Here is a clear example: "A revolutionary transformation of social life will not simply happen through piecemeal change. Rather, we believe it will occur only as the result of a prolonged struggle based on hope and a total vision of a qualitatively new society, waged by those social classes and groups who stand to benefit from the new era" (p. 17).

By placing great faith in an impending revolution, they were able to disparage incremental reforms. In this line of thought, policies either promote a socialist revolution or else "reproduce" capitalism (van den Berg 1988). This "either/or" reasoning allows profound historical changes such as the huge expansion of public education to be not only dismissed as mere tinkering with the system but even regarded as a sinister strategy to hinder the unfolding of a socialist consciousness that would otherwise emerge. Such an allegation makes sense only by referring to how some imagined nonincorporated working class would act on their objective interests. Yet, other than airily alluding to some far-projected surge in class consciousness, reproduction theorists offered few specifics. For example, Bowles and Gintis (1976, pp. 14–15) could submit only this: "While there is nothing inevitable about a democratic socialist movement bursting through the fabric of irrational social relationships, the possibility grows yearly. . . . The economic transformation which we envision and which is the basis for our optimism, is so far-reaching and



total in its impact on social life as to betoken a new stage in the development of U.S. society. Moreover, it requires a historical consciousness on the part of citizens of a type uncommon in our history." This faith in revolution cannot be regarded as empirically based, since Bowles and Gintis mustered no compelling reasons for why they believed the potential for revolution to be growing in the mid-1970s. Paradigm boundaries thus originated from their tacit adoption of an immanent philosophy of history.

But reproductionists faced the difficult task of retaining this rigorous critique of capitalism while accommodating the nonrevolutionary character of the working class. Since they were chiefly concerned with detailing how the capitalist system ensures its seemingly uncertain continuation by keeping contradictory forces at bay, they were left with few concrete hints of where to look for change. Reproductionism thus fell short of supplying any tangible indication for potential socialist transformation (Giroux 1983, p. 259). Resistance theory was proclaimed as a solution, as it retained this critique of capitalist schooling but now specified the agents who will struggle against it. Thus, the movement from reproduction theory to resistance theory can be clearly seen as an attempt to resolve the tensions within immanentism. Reproduction theory emphasizes how capitalism debases the working class, and resistance theory embodies the class struggle, the motor of change. Yet these questions remain: Given the ostensible similarities between resistance theory and subcultural theory, from where does resistance theory's politicized interpretation emerge? And in what ways does it reinforce paradigm boundaries?

Resistance theory represents a particular application of cultural Marxism that is tailored for youth. While high school dropouts may appear to non-Marxists as unlikely agents of revolution, these youth fit well into the cultural Marxist images of immanent change in the late 1970s and early 1980s. Resistance studies championed adversarial student behaviors when they were deemed compatible with the classic conception of the "making" of the working class, such as those that hint at solidarity, antagonism to authority, and so on. They equated student subcultures with "proletarian resistance," implicitly assuming that teenage antics necessarily reflect class experiences rather than common adolescent pre-occupations with fashion, excitement, and peer group solidarity (Frith 1985), an equation that makes sense only if one associates any type of rebellion with an anticapitalist impulse. In contrast, when these students are not oppositional, they are regarded as passively lacking "true" insight, presumably due to a lack of access to the authentic proletarian culture that rebellious youth enjoy (e.g., Willis 1977; Corrigan 1979). But an unacknowledged irony of applying the class formation model to

analyses of schools is its striking resemblance to the discredited cultural deprivation approach. Resistance theorists inadvertently redeploy the deprivationists' assumption that the working class is truly distinct from the middle class, as opposed to merely adapting to unequal opportunities (i.e., that they really are less competitive and less striving for upward mobility). By not acknowledging these similarities, resistance theorists avoid the indictments that were hurled at the deprivationists. But in the shadow of accusations of "blaming the victim," why do they revive these previously discredited ideas?

The answer lies in the tension caused by having to identify both the forms of capitalist oppression and agents of transformation. While reproductionism made for effective criticism, its reduction of all cultural distinctions to differential placement in opportunity structures had an unsettling implication. If the theory was correct, any future equalization of educational opportunities within capitalism would slowly remove cultural distinctions. And if everyone eventually shares a middle-class culture, then no unincorporated agent of change remains. This disconcerting specter motivated resistance theorists to rework cultural deficit theory, because they wished to tacitly assert that the working class remains, to some degree, culturally distinct.

Despite this similarity to cultural deprivation theory, resistance theory nevertheless differed because of its key premise: disadvantaged class origins and educational failure facilitate, at some level, leftist inclinations. Yet this assumption forces resistance theory into a series of recourses that have no counterpart in the mainstream.

Consider the difficulties that arise when one attempts to square this premise with empirical data. Little evidence suggests that working-class school rejection is associated with leftist political participation and activism. For instance, Downes (1966) prophetically warned that British early leavers were possible fodder for protofascist movements, foreshadowing British riots in the late 1970s. Since then, most research documents that, historically, middle-class students have been more political than working-class youth (Frith 1985). Case studies suggest that typical dropouts bear little resemblance to the archetypal student activists of the 1960s, in terms of the types of schools they attend, their degree of integration into school organizations, and their social backgrounds. The profiles of leftist students are typified by relatively privileged socioeconomic backgrounds, attendance at elite universities, solid grades, and active involvement in school activities (Levitt 1984; Tanner 1988; Paulsen 1991). In terms of gender politics, low-achieving and delinquent females are more traditional than academically striving young women and are less supportive of the women's movement (Steffensmeir et al. 1989; Holland and Eisenhart 1990; Davies 1993). The paradoxical conclusion to be drawn is that

the relation between radicalism, social class, and school disaffection is largely *opposite* to that premised in resistance theory. Since most politicized students hail from the middle class, resistance theorists' linking of school failure to leftist activism appears to invert the empirical relation between student type and political stance.

Perhaps anticipating these criticisms, resistance arguments tend to be heavily qualified, most notably with the disclaimer that much resistance is symbolic (CCCS 1981; Giroux 1983). Though Willis (1983) admits that much resistance is experienced as plain "fun and excitement," he nevertheless argues that this fun and excitement articulates an opposition to capitalist schooling. In the absence of overt radical opposition among working-class youth, resistance theorists react with the caveat that radical insights are not necessarily consciously held nor intended but are symbolized in distinctive class mores and styles, as decoded by the analyst. For instance, Aronowitz and Giroux (1993, p. 101) offer this disclaimer: "We want to stress that the theoretical construct of resistance rejects the positivist notion that the meaning of behavior is synonymous with a literal reading based on immediate action. Instead resistance must be viewed from a theoretical starting point that links the display of behavior to the interests that underlie its often hidden logic, a logic that also must be interpreted through the historical and cultural mediations that shape it."

This passage provides an analytic rationale for downplaying an actors' own intentions and meanings in light of some theoretically imputed interest. While this may appear as a nuanced argument to some, others may read it as a fallback plan for dealing with discomforting evidence. To encourage empirical observation yet then reject "literal readings" when necessary allows one to avoid data that fail to meet the expectations of the theory. Under the banner of antipositivism, Aronowitz and Giroux thus prescribe a rationale that, by favoring "interests" over actors' intentions and immediate context, ironically nullifies the strength of ethnographic methods vis-à-vis survey research—its sensitivity to subjective meanings and the surrounding context.

Such caveats serve an important role within resistance theory, because empirical examples offered in school ethnographies are either contradictory or unconvincing. For instance, one of resistance theory's founding claims is that working-class youth subcultures express a solidarity that opposes middle-class competitiveness. Some simply present this idea as an unquestioned given (e.g., Aronowitz and Giroux 1993, p. 102). Those who attempt to establish this claim empirically offer thin support. For instance, in Willis's study the so-called solidarity of rebellious lads is most obviously highlighted in their out-group rejection of effeminate males, females, and racial minorities—hardly a celebration of class solidarity.

Other examples risk trivializing the concept. Eckert's (1989) sole evidence for her claims of greater egalitarian social relations among suburban Detroit working-class youth were their more age-heterogeneous friendship groups, their exchange of cigarettes, and their stealing from one another (which Eckert admits is a rather convoluted form of solidarity). Similar problems can be seen in the field research of Corrigan (1979) and McLaren (1986). But besides being unconvincing, these researchers fail to confront what they inadvertently document: these peer groups themselves are hierarchical and competitive along nonacademic criteria, such as their fighting or sexual prowess. If such subcultures reject middle-class competitiveness in favor of egalitarianism, then why do they have these traits and how will they change?

Even if one sets these questions aside and accepts the "symbolic resistance" qualifier, one must note that there is little spectacular data to work with; many researchers resort to designating "resistance" as almost anything ranging from drinking and fighting with teachers to less conspicuous phenomena such as boredom, indifference, laziness, and quiet behavior (e.g., Anyon 1981; Everhart 1983). In fact, almost anything short of joyful compliance has been infused with partisan meaning. Thus, though the qualified argument breathes new life into the analysis, it comes at the cost of diluting the substance of resistance. When such a range of behaviors can represent the concept, it is obliged to accommodate virtually anything and becomes highly ambiguous.

Part of this problem may exist because much school rejection is transient and ephemeral. As Willis and others document, many early leavers later regret their decision to leave school, having recognized their ensuing lower life chances, and thus retract many (what the theorists consider to be) "class insights" *after* gaining labor market experiences. That is, they endorse the "school as mobility" ethic after having encountered the harsh realities of work or unemployment. These ephemeral qualities of school rebellion have been accommodated within resistance theory, however. A noted disclaimer is that youth subsequently discard these insights because adult roles such as employment and parenting impose various hegemonic or conservatizing effects (Willis 1977; Aronowitz and Giroux 1985). However, this qualifier forces the theory onto unsteady ground, for it effectively inverts a central resistance tenet: working-class insights are rooted in the experience of wage labor, and young rebels reject school in part by drawing upon the worldly knowledge of their working relatives. Thus, this recourse logically implies the absurdity that critical penetration into the character of capitalism reaches its zenith among working people at age 16 or 17.

Since the mid-1980s, definitions of resistance have again shifted, as post-Marxists have grown skeptical of the political potential of most

"symbolic resistance," especially the more reactionary elements of counterschool subcultures. Ethnographers clearly documented that male pranksters were often virulently sexist and racist and that many female dropouts were, in the final analysis, gender traditionalists. Resistance theorists such as Willis recognized these facts, of course, but took a "wait and see" stance on the hope that, with further development and perhaps some critical pedagogy, the reactionary components could be expunged or rechanneled into a purer form (e.g., Apple 1985). But post-Marxists insist that not all oppositional behavior is to be considered resistance. Their preference has been to expunge the category of all that is corrupted by sexism, racism, and individualism. Consequently, post-Marxist definitions of "true" resistance includes only opposition that possesses "authentic emancipatory promise," meaning that it incorporates leftist struggles, such as feminism, antiracism, and gay liberation (Giroux 1989; Sultana 1989; Fernandes 1988; Aronowitz and Giroux 1993).

This shift to a more demanding criterion of resistance has had two important effects. First, as noted above, it has led American post-Marxists to withhold their approval of working-class males. Instead of deciphering school pranks as signals of nascent radicalism, post-Marxists have grown increasingly wary of these youths' supposed susceptibility to neoconservative populism. Indeed, in a remarkably brief span of time (1977–90) working-class males went from being hailed as a possible radical vanguard (e.g., Willis 1977; Corrigan 1979) to being distrusted as potential supporters of the New Right backlash (e.g., Apple 1989; Weis 1990). This new assessment is not empirically based, since the actions of working-class youth themselves have scarcely changed over this time period. Rather, it reflects the shifting premises of critical approaches.

Second, and more important, few disaffected youth actually meet the standard of simultaneous progressive stances—being anticapitalist, anti-patriarchal, antihomophobic, and promulticulturalist—only the combination of which is held to be true resistance. Indeed, among Sultana (1989), Fernandes (1988), Giroux (1983), Giroux and McLaren (1989), McLaren (1989), and Aronowitz and Giroux (1985, 1991, 1993), *no* examples of authentic resistance among students are offered. In fact, the lone documented instance of actors engaging in true resistance are the critical pedagogues themselves (Fernandes 1988). Consequently, the more exacting standard of resistance entails a key sacrifice. Without announcement, it transforms "resistance" from an empirical category into an unachieved ideal whose political purity comes at the cost of a retreat from the messy reality of life in schools. It squeezes out virtually all empirical referents of resistance, making the category largely bereft of observable content.

To review, resistance theorists invoke an number of original caveats; they at times avoid "literal readings" of behavior, stress the symbolic

aspects of student deviance, invoke either highly diluted or rigorous standards of resistance, and offer abstract analyses of cultural hegemony when former rebels mellow with age. The point of documenting this series of recourses is that they mark paradigmatic boundaries. Resistance theory is remarkably similar to subcultural theory (and in some respects, cultural deprivation theory) until it is forced to deal with empirical challenges to its core presumption, namely that working-class youth should be oppositional toward school and politically leftist. When working-class youth empirically fail to meet the latter expectation—which is usually the case—the theory resorts to either a diluted version of “resistance” or a more rigorous but empirically empty definition of the concept, in both cases reverting to rather wobbly reasoning in order to safeguard the original premise. Only then does resistance theory become unlike the mainstream. Its ambivalence toward empirical evidence sets in motion a series of caveats that lacks a corollary in the mainstream.

Critical theory could move in the direction of the mainstream if some of these qualifications were tested. One could investigate whether, for instance, unmarried and childless former dropouts are more radical than other dropouts, or attempt to measure some of the hypothesized hegemonic effects, such as degree of internalization of the mobility ethic, or see whether only certain types of working-class youth have leftist leanings. But this type of investigation is not what occurs. Rather, critical theorists frame their qualifications to make their domain assumption—that capitalist schooling is unbearable to the working class and necessarily evokes a leftist response from these youth—immune to damaging evidence. Announcing that much of this resistance is actually symbolic and thus should not be read literally is perhaps the most graphic illustration of this drift into nonfalsifiability.

However, these practices—the romanticization of the working class, the implicit presumption that history moves in necessary directions, the reduction of most social phenomena to class struggle, and the attributing of “objective interests” to groups—have all been renounced by post-Marxists (see Laclau and Mouffe [1987] for a post-Marxist manifesto). But is post-Marxism able to evade the problems that I have identified as endemic to critical theory?

Certainly post-Marxists do not see radical social change to be near on the horizon. In part a response to 1980s and 1990s neoconservatism, post-Marxists are cautious in their estimations of radical social change and in their judgments of educational reforms (see Wexler 1987; Apple 1989; Aronowitz and Giroux 1985, 1991; Giroux and McLaren 1989; Weis 1990). While the economic buoyancy of the 1960s and 1970s allowed radicals to negatively portray the existing school reforms as actually hindering a spontaneous mass movement toward more profound social

change, the fact that some of these reforms have since been reversed holds that reasoning to question. Radicals can now ill afford to dismiss liberal reforms outright. As loans, accessibility measures, and progressive curricula are jeopardized, they are less easily seen as co-optations of an otherwise primed revolutionary movement, as assumed in reproductionist logic. Post-Marxists now reevaluate some of the past Marxist criticisms of liberal schooling and employ a more gradualist model of educational change (e.g., Aronowitz and Giroux 1985, 1993).

This new conception of change leads to a rejection of the cultural Marxist model used in resistance theory, which implicitly designated informal groups and the development of a grassroots counterculture to be the foundation of a potential radical uprising. Since 15 years of ethnography uncovered countless instances of "submerged" opposition yet few inspiring tangible examples, post-Marxists doubt the political potential of overly individualized forms of resistance. For instance, Carnoy (1989, pp. 20–21) notes in a recent discussion of school troublemaking: "A number of analysts have called such conflict 'resistance.' . . . however . . . *exit*—acting out against or dropping out from the school—without political engagement is not resistance. Rather, it is essentially a personal attempt to escape from a costly, painful situation. And the only way that such exit becomes political is when the larger community—either the dominant business class or subordinate groups themselves—make school frustrations and dropouts a political issue." Carnoy's reasoning indicates not only a disenchantment with the limitations of informal resistance but also how the search for agents has turned full circle. As critical theorists find informal domains and unorganized acts of diffuse rebellion increasingly deficient, even dismissing them as "romantic individualism" (Wexler 1987), they are returning to formal channels. In their place, post-Marxists look for more articulate and organized groups such as sexual minorities, African-Americans, and Native Americans, who, being oppressed, are presumably motivated not merely to reform the education system through better representation and access but ultimately to *transform* it. Post-Marxists assume that these suppressed voices will express more than outrage against exclusion, limited opportunities, and unequal chances. Rather, these groups are seen to embody a dormant yet massive critique of schools' complicity in domination (Aronowitz and Giroux 1985).

But to view agents as possessing a higher, yet hidden, potential is familiar. It parallels Marxist treatments of unions, which Marxists nominated not only to fight for workers' immediate interests but to be vehicles of revolution. Thus, post-Marxists face the old question of how ostensibly reformist movements will eventually shift their aims to overturning the system (Lockwood 1981). Such groups can appear radical precisely because their exclusion has hitherto prevented them from developing a

pattern of compromise. However, their course of development may change when they grow stronger, for additional strength may temper their tactics. For instance, recent trends of blacks pressing for an Afrocentric curriculum are celebrated by post-Marxists as having potential for linkages to other radical causes. But they ignore how these demands often appear in tandem with African-American entrepreneurship and economic self-reliance, phenomena that are more easily interpreted in a liberal pluralist model than a grander design of radical democratic socialism.

When facing this problem, post-Marxists do not invoke immanentist claims of lawlike movements of history, as did their Marxist predecessors. Nevertheless, post-Marxists espouse a problematic theory of what drives social change. For them, agency is charged by "desire" and the "imaginary," that is, proclivities toward creating an alternative world as opposed to what currently exists (Aronowitz and Giroux 1985, p. 18). While the education system is seen to suppress and co-opt this imaginary, this same imaginary is also seen as the basis for opposition. Accordingly, the proclaimed task of critical education is to instill a "critical literacy" to empower subordinate groups and help them recover their "true selves" (Giroux and McLaren 1989). But to invoke notions like "desire" and the "imaginary" echoes the cavalier Marxist assumption that designated agents "naturally" gravitate toward socialism, or in this case, "emancipation." Further, by professing their goal as helping groups "recover themselves" and reject "repressed and distorted identities" that schools impose (Aronowitz and Giroux 1985, p. 137), post-Marxists presume that, given the tools for critical reflection, masses of disadvantaged youth would undoubtedly share the vision of critical theory. For these theorists, it is unimaginable that many of these youth may hold different understandings of their ultimate interests and hence may be less enthused about the radical causes they are being urged to struggle for. This problem appears in a recent theoretical discussion that *ends* with the very question with which it should commence (Giroux and McLaren 1989, p. 251): What if the critical teachers' agenda differs from what their students and families want? By not supplying an answer, they leave unsubstantiated their fundamental rationale: that only critical theorists know what ordinary folk really desire.

More exactly, this rationale is reminiscent of the Marxist assumption that these groups do not know their own true interests. For instance, when students from various designated groups do not act in favor of alternative schooling, post-Marxists deem this failure to act the result of cultural barriers. Aronowitz and Giroux (1985, p. 173), for example, note that "it must be admitted that the demand for a new curriculum does not correspond currently to the economic and cultural perceptions of the majority of students, especially minorities and women." They attribute this situ-



ation to the predominance of the "equality of opportunity" ideology in society at the expense of more radical discourses. This explanation reenacts the familiar Marxist recourse to explain the failure of designated groups to act in their "real" interests as the consequence of ideological obfuscation.

Post-Marxists thus leave intact the implicit Marxist practice of ascribing radical interests to a collective agent and differ only by surreptitiously *extending* this scheme to women and various minorities. They replace the imputed working-class interest in socialism and its subterranean narrative of a spontaneously socialist-bound proletariat by imputing several groups with interests in radical democracy with a pretext that this interest represents ultimate human yearnings for emancipation. Besides implying that post-Marxists are somehow the exclusive confidantes of primordial human desires, this claim repeats the Marxist practice of projecting their own political aspirations—in this case, the goals and rhetoric of coalition-building for radical democracy—as the objective interests of disadvantaged groups. Old dilemmas remain: how to explain how movements that are ostensibly reformist will turn more radical and how to account for the failure of designated groups to recognize their interests. Thus, the discussion of "desire" and "emancipation" serves to restate the Marxist philosophy of history in a more up-to-date language.

#### SUMMARY AND CONCLUSION

This article has highlighted a rift between two paradigms in the sociology of education. Two primary arguments challenge critical theorists' assessments of this division.

First, I argue that this split is not rooted in differences in empirical technique, in praxis, or in an incompatibility between positivism and interpretivism, since mainstream and critical researchers alike have used multiple empirical methods and have had their work informed by policy and political concerns. Moreover, the theories have some strikingly similar substantive content. Resistance theory and subcultural theory reached largely parallel empirical judgments, as did reproduction theory and the early status attainment research. These theories differed only in interpretation. The source of this difference is a set of key "immanentist" presumptions that instill an ambivalence to empirical data. The implicit presumptions of a necessary movement of history and of the working class's objective interest in socialism allowed reproduction theorists to assert that the educational inequalities were inevitable in capitalism and that reforms were thus futile. This assumption, since discarded, encouraged the reproductionists to define themselves in opposition to a more cautious "liberal" mainstream.

Resistance theory diverged from subcultural theory only when con-

fronted with data that did not support its expectations. This crisis produced a series of theoretical caveats that have no counterpart in the mainstream. Resistance theorists held the unshakable presumption that subordinate youth are inherently oppositional to capitalist schooling and that this antagonism does not just express a desire for better educational opportunities but ultimately represents a yearning for more fundamental change. Finally, though post-Marxists claim to have avoided these assumptions, they continue to implicitly attribute objective interests to subordinate groups and assume that these groups have a natural affinity with radical social causes. These assumptions continue to undergird the boundary between critical and mainstream sociology of education.

Second, I outline how critical theorists' images of working-class students have changed markedly over the past two decades. After reproductionists portrayed these youth as passive victims of capitalism, resistance theorists transformed this image into that of a cultural terrorist. Post-Marxists now offer a mixed portrait of either politically pure world citizens or potential traitors who may join the New Right. I argue that these shifts reflect how critical theorists' conceptions of radical agents and modes of transformation have fluctuated. These shifts are set in motion when designated groups fail to fulfill their deemed potential. The movement from the reproductionist all-or-nothing call for revolution to resistance theory's cultural Marxism to post-Marxism's new social movements reflects substituted estimations of how and where radical change will occur, rather than any empirical changes among students or advances in knowledge about schooling.

Perhaps these conclusions can be a cause for optimism among those who wish greater communication across paradigm lines. Mainstream and critical theories converge in many ways and often diverge only as a consequence of assumptions that I have labeled "immanentist." These assumptions are, I argue, substantive claims that can be held or discarded by researchers regardless of their epistemological or methodological preferences. They are more accurately seen as leaps of faith without a necessary connection to any methodology. In practice, the type of empirical method we choose, or whether our work is practically oriented or "realist" or "interpretivist," does not completely determine our conclusions. And if most types of research are not inherently incompatible, then there is greater room for dialogue and synthesis than some currently imagine.

## REFERENCES

- Aggleton, Peter, and Geoff Whitty. 1985. "Rebels without a Cause? Socialization and Subcultural Style among the Children of the New Middle Class." *Sociology of Education* 58 (1): 60-72.

- Alexander, Jeffrey. 1982. *Theoretical Logic in Sociology*. Vol. 1, *Positivism, Presuppositions, and Current Controversies*. Berkeley and Los Angeles: University of California Press.
- Anyon, Jean. 1981. "Elementary Schooling and Distinctions of Social Class." *Interchange* 12 (2-3) 118-32.
- Apple, Michael. 1979. *Ideology and Curriculum*. London: Routledge & Kegan Paul.
- . 1985. *Education and Power*. Boston: Ark.
- . 1989. "The Politics of Common Sense: Schooling, Populism and the New Right." Pp. 32-49 in *Critical Pedagogy, the State, and Cultural Struggle*, edited by Henry Giroux and Peter McLaren. Albany: State University of New York Press.
- Aronowitz, Stanley. 1981. Preface to the Morningside Edition. Pp. ix-xiii in *Learning to Labour: How Working Class Kids Get Working Class Jobs*, by Paul Willis. New York: Columbia University Press.
- Aronowitz, Stanley, and Henry Giroux. 1985. *Education under Siege: The Conservative, Liberal, and Radical Debate over Schooling*. South Hadley, Mass.: Bergin & Garvey.
- . 1991. *Postmodern Education*. Minneapolis: University of Minnesota Press.
- . 1993. *Education Still under Siege*, 2d ed. Toronto: OISE Press.
- Banks, Olive. 1976. *The Sociology of Education*. New York: Schocken.
- Blau, Peter, and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in Capitalist America*. New York: Basic.
- Brake, Michael. 1985. *Comparative Youth Culture. The Sociology of Youth Subcultures in America, Britain, and Canada*. London: Routledge & Kegan Paul.
- Campbell, Richard T. 1983. "Status Attainment Research: End of the Beginning or Beginning of the End?" *Sociology of Education* 56:47-62.
- Carnoy, Martin. 1989. "Education, State, and Culture in American Society." Pp. 3-23 in *Critical Pedagogy, the State, and Cultural Struggle*, edited by Henry Giroux and Peter McLaren. Albany: State University of New York Press.
- CCCS (Center for Contemporary Cultural Studies). 1981. *Unpopular Education: Schooling and Social Democracy in England since 1944*. London: Hutchinson.
- Cohen, Albert. 1955. *Delinquent Boys: The Culture of the Gang*. New York: Free Press.
- Cohen, Stanley. 1980. *Folk Devils and Moral Panics*. Oxford: Martin Robinson.
- Coleman, James, and Torsten Husen. 1985. *Becoming Adult in a Changing Society*. Paris: OECD.
- Collins, Randall. 1986. "Is 1980s Sociology in the Doldrums?" *American Journal of Sociology* 91 (6): 1336-55.
- . 1989. "Proscience or Antiscience?" *American Sociological Review* 54 (1): 129-39.
- Connell, Robert, D. Ashendon, D. Kessler, and G. Dowsett. 1982. *Making the Difference*. Sydney: George Allen & Unwin.
- Corrigan, Philip. 1979. *Schooling the Smash Street Kids*. London: Macmillan.
- Crouch, Colin. 1982. *Trade Unions: The Logic of Collective Action*. Glasgow: Fontana.
- Davies, Scott. 1993. "Exploring Reproduction through Resistance in Ontario High Schools." Paper presented at the annual meetings of the American Sociological Association, Miami.
- . 1994a. "Class Dismissed? Student Opposition in Ontario." *Canadian Review of Sociology and Anthropology*. 31 (3): 421-44.
- . 1994b. "In Search of Rebellion and Resistance among High School Drop-outs." *Canadian Journal of Sociology* 19 (3): 331-50.

- Downes, David. 1966. *The Delinquent Solution: A Study in Subcultural Theory*. London: Routledge & Kegan Paul.
- Downes, David, and Paul Rock. 1988. *Understanding Deviance: A Guide to the Sociology of Crime and Rule Breaking*, 2d ed. Oxford: Clarendon.
- Eckert, Penelope. 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teachers College Press.
- Everhart, Robert. 1983. *Reading, Writing, and Resistance*. Boston: Routledge & Kegan Paul.
- Fernandes, Joao V. 1988. "From the Theories of Social and Cultural Reproduction to the Theory of Resistance." *British Journal of Sociology of Education* 9 (2): 169-80.
- Frith, Simon. 1985. "The Sociology of Youth." Pp. 303-68 in *Sociology: New Directions*, edited by M. Haralambos. Ormskirk: Causeway.
- Gaskell, Jane. 1985. "Course Enrollment in the High School: The Perspective of Working-Class Females." *Sociology of Education* 58:48-59.
- Giroux, Henry. 1983. *Theory and Resistance in Education: A Pedagogy for the Opposition*. South Hadley, Mass.: Bergin & Garvey.
- . 1989. "Schooling as a Form of Cultural Politics: Toward a Pedagogy of and for Difference." Pp. 125-51 in *Critical Pedagogy, the State, and Cultural Struggle*, edited by Henry Giroux and Peter McLaren. Albany: State University of New York Press.
- Giroux, Henry, and Peter McLaren, eds. 1989. *Critical Pedagogy, the State, and Cultural Struggle*. Albany: State University of New York Press.
- Goldthorpe, John. 1988. "Intellectuals and the Working Class in Modern Britain." In *Social Stratification and Economic Change*, edited by David Rose. London: Hutchinson.
- Gouldner, Alvin. 1985. *Against Fragmentation: The Origins of Marxism and the Sociology of Intellectuals*. New York: Oxford University Press.
- Halsey A. H., Anthony F. Heath, and John Ridge. 1980. *Origins and Destinations: Family, Class, and Education in Modern Britain*. Oxford: Clarendon.
- Hammersley, Martyn. 1985. "From Ethnography to Theory: A Programme and Paradigm in the Sociology of Education." *Sociology* 19 (2): 244-59.
- Hargreaves, Andy. 1982. "Resistance and Relative Autonomy Theories: Problems of Distortion and Incoherence in Recent Marxist Analyses of Education." *British Journal of Sociology of Education* 3 (2): 107-26.
- Hargreaves, David. 1967. *Social Relations in a Secondary School*. London: Routledge & Kegan Paul.
- Harp, John. 1991. "Political Economy/Cultural Studies: Exploring Points of Convergence." *Canadian Review of Sociology and Anthropology* 28 (2): 206-24.
- Holland, Dorothy, and Margaret Eisenhart. 1990. *Educated in Romance*. Chicago: University of Chicago Press.
- Hyman, Herbert. 1953. "The Value Systems of Different Classes: A Social Psychological Contribution to the Analysis of Stratification." Pp. 426-42 in *Class Status and Power*, edited by Reinhard Bendix and Seymour Martin Lipset. Glencoe, Ill.: Free Press.
- Karabel, Jerome, and A. H. Halsey. 1977. "Education Research: A Review and an Interpretation." Pp. 1-85 in *Power and Ideology in Education*, edited by Jerome Karabel and A. H. Halsey. New York: Oxford University Press.
- Lacey, Colin. 1970. *Hightown Grammar*. Manchester: University of Manchester Press.
- Laclau, Ernesto, and Chantal Mouffe. 1987. "Post Marxism without Apologies." *New Left Review* 166:79-106.
- Leong, Laurence Wei-Teng. 1992. "Cultural Resistance: The Cultural Terrorism of British Male Working-Class Youth." *Current Perspectives in Social Theory* 12: 29-58.

- Levitt, Cyril. 1984. *The Children of Privilege*. Toronto: University of Toronto Press.
- Liston, Daniel. 1988. *Capitalist Schools: Explanations and Ethics in Radical Studies of Schooling*. New York: Routledge.
- Livingstone, David. 1983. *Class Ideologies and Educational Futures*. New York: Falmer Press.
- Lockwood, David. 1981. "The Weakest Link in the Chain? Some Comments on the Marxist Theory of Action." *Research in the Sociology of Work* 1:435-81.
- Macleod, Jay. 1987. *Ain't No Makin' It: Levelled Aspirations in a Low-Income Neighbourhood*. Boulder, Colo.: Westview.
- Marshall, Gordon. 1983. "Some Remarks on the Study of Working-Class Consciousness." *Politics and Society* 12 (3): 263-301.
- McLaren, Peter. 1986. *Schooling as a Ritual Performance*. London: Routledge & Kegan Paul.
- . 1989. *Life in Schools: An Introduction to Critical Pedagogy in the Foundations of Education*. White Plains, N.Y.: Longman.
- McRobbie, Angela. 1981 "Settling Accounts with Subcultures: A Feminist Critique." Pp. 111-24 in *Culture, Ideology, and Social Process*, edited by Tony Bennett, Graham Martin, Colin Mercer, and Janet Woollacott. London: Open University Press.
- Mehan, Hugh. 1992. "Understanding Inequality in Schools. The Contribution of Interpretive Studies." *Sociology of Education* 65 (1). 1-20.
- Morrow, Raymond. 1985. "Critical Theory and Critical Sociology." *Canadian Review of Sociology and Anthropology* 22 (5): 710-45.
- Murdock, Graham, and Guy Phelps. 1973. *Mass Media and the Secondary School*. London: Macmillan.
- Parkin, Frank. 1979. *Marxism and Class Theory: A Bourgeois Critique*. London: Tavistock.
- Paulsen, Ronnelle. 1991. "Education, Social Class, and Participation in Collective Action." *Sociology of Education* 64 (2): 96-110.
- Rist, Ray. [1977] 1990. "On the Relations among Educational Research Paradigms: From Disdain to Detente." Pp. 81-95 in *Education and Society: A Reader*, edited by Kevin Dougherty and F. Hammock. San Diego: Harcourt Brace Jovanovich.
- Ryan, William. 1971. *Blaming the Victim*. New York: Vintage.
- Sewell, William, and Robert Hauser. 1980. "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." *Research in Sociology of Education and Socialisation* 1:59-99.
- Steffensmeir, Darrell J., Emilie A. Allan, Miles D. Harer, and Cathy Streifel. 1989. "Age and the Distribution of Crime." *American Journal of Sociology* 94 (4): 803-31.
- Stinchcombe, Arthur. 1964. *Rebellion in a High School*. Chicago: Quadrangle Books.
- Sultana, Ronald G. 1989. "Transition Education, Student Contestation, and the Production of Meaning: Possibilities and Limitations of Resistance Theories." *British Journal of Sociology of Education* 10 (3): 287-309.
- Tanner, Julian. 1988. "Youthful Deviance." Pp. 323-59 in *Deviance: Conformity and Control in Canadian Society*, edited by Vince Sacco. Scarborough: Prentice-Hall Canada.
- . 1990. "Reluctant Rebels: A Case Study of Edmonton High School Dropouts." *Canadian Review of Sociology and Anthropology* 27 (1): 74-94.
- Tanner, Julian, Scott Davies, and Bill O'Grady. 1992. "Immanence Changes Everything: A Critical Comment on Labour Process Theory and Class Consciousness." *Sociology* 26 (3): 439-53.
- Valentine, Charles. 1968. *Culture and Poverty: Critique and Counter Proposals*. Chicago: University of Chicago Press.
- van den Berg, Axel. 1988. *The Immanent Utopia: From Marxism on the State to the State of Marxism*. Princeton, N.J.: Princeton University Press.

- . 1993. "Creeping Embourgeoisement? Some Comments on the Marxist Discovery of the New Middle Class." *Research in Social Stratification and Mobility* 12:295–328.
- Walker, J. C. 1986. "Romanticizing Resistance, Romanticizing Culture: Problems in Willis' Theory of Cultural Production." *British Journal of Sociology of Education* 7 (1): 59–80.
- Wallerstein, Immanuel. 1986. "Marxisms as Utopias: Evolving Ideologies." *American Journal of Sociology* 91 (6): 1295–1308.
- Weis, Lois. 1990. *Working Class without Work: High School Students in a De-Industrializing Economy*. New York: Routledge.
- Wexler, Philip. 1987. *Social Analysis of Education: After the New Sociology*. London: Routledge & Kegan Paul.
- Willis, Paul. 1977. *Learning to Labour: How Working Class Kids Get Working Class Jobs*. New York: Columbia University Press.
- . 1983. "Cultural Production and Theories of Production." Pp. 107–138 in *Race, Class, and Education*, edited by Len Barton and S. Walker. London: Crook Helm.
- . 1990. *Common Culture: Symbolic Work at Play in the Everyday Cultures of the Young*. Boulder, Colo.: Westview.
- Woods, Paul. 1983. *Sociology and the School: An Interactionist Viewpoint*. London: Routledge & Kegan Paul.

# Strikebreaking or Solidarity in the Great Steel Strike of 1919: A Split Labor Market, Game-Theoretic, and QCA Analysis<sup>1</sup>

Cliff Brown and Terry Boswell  
*Emory University*

Split labor market theory provides insight into the development of ethnic and racial antagonism but has failed to address interracial solidarity and has tended to ignore the role of the state. The authors modify Heckathorn's formal model of collective action to derive predictions concerning the possibility of interracial solidarity or strikebreaking given split labor market conditions. Predictions are then examined using qualitative comparative analysis for 16 northern cities that participated in the 1919 steel strike. Results show that interracial solidarity developed in cities that had strong local union organizations and nonrepressive governments, while black strikebreaking emerged in cities with higher proportions of recent black migrants and either repressive local governments or weak unions.

## INTRODUCTION

Theories of racial antagonism frequently fail to explain why and when people overcome an environment of racism to achieve interracial solidarity. This lacuna is particularly evident in split labor market theory (Bonacich 1972; Wilson 1978; Marks 1981; Boswell 1986). Racial conflict is generated, according to the theory, when labor market competition is visibly split between higher paid workers of the dominant ethnicity and low-cost, usually migrant or immigrant, minority labor. Minority strikebreaking and dominant discrimination are prevalent when minority workers are disproportionately short-term sojourners who would not ben-

<sup>1</sup> An earlier version of this paper was presented at the meetings of the American Sociological Association, Pittsburgh, August 1992. The authors wish to thank Alex Hicks, Doug Heckathorn, Frank Lechner, Rick Robinson, John Brueggemann, Joya Misra, Steve Harvey, Tim Brezina and the *AJS* reviewers for their helpful comments. Direct correspondence to Cliff Brown, Department of Sociology, Emory University, Atlanta, Georgia 30322.

efit from long union struggles. But does the logic of the theory apply in the converse? Specifically, is interracial solidarity most likely when few minority workers are sojourners? Or, is sojourning secondary to factors emphasized in other theories, such as the sanctioning power of unions (Hechter 1987) or class articulation of the state (Burawoy 1981)?

To address these and related questions, we examine the pattern of interracial strikebreaking and solidarity that accompanied the industry-wide steel strike of 1919, a watershed in race relations and labor history. The interwar period is critical for the development of split labor market theory as well, because strikebreaking and riots during the labor disturbances of the 1910s and 1920s provide key illustrative cases (see Bonacich 1976; Wilson 1978, pp. 71–77). In the theory's development, these cases served as a "crucial test" (Eckstein 1975). The specified conditions are so obviously present and the events are of such historical importance that the absence of expected results would invalidate the theory. Had strikebreaking and antagonism not occurred, split labor market postulates would be highly suspect.

We return to the events of 1919 in order to test the theory's applicability to solidarity as well as strikebreaking, focusing on 16 communities in which the specific strike events took place. We also examine whether the theoretical centrality of sojourning migration withstands comparison to alternative community factors, specifically the historical strength of the labor movement and the impact of state repression. In the following sections we first discuss the limitations of split labor market theory, then we use a formal theory of collective action (Heckathorn 1990a; 1990b; 1991) to model the union mobilization of minority workers and dominant workers given different split labor market conditions.<sup>2</sup> This is followed by a historical overview of race and labor relations during the period and a systematic comparison of cases using qualitative comparative analysis (QCA; Ragin 1987; Drass and Ragin 1986). Finally, we discuss the implications of our findings for building split labor market theory and of our methods for use in historical sociology.

#### SPLIT LABOR MARKETS AND SOLIDARITY

The dynamics of split labor markets have been explored in a number of important studies (Bonacich 1972, 1975, 1976, 1979; Wilson 1978; Marks

<sup>2</sup> We use the term "dominant" throughout to refer to the numerically larger and more powerful racial group (in this case, white workers). Thus, the term "dominant union" reflects the fact that labor organizations of the period were largely controlled by and composed of white labor.



1981; Boswell 1986). A split labor market exists when labor costs of equally skilled workers differ in terms of race or ethnicity (Bonacich 1972, p. 549). Lower labor costs for minority groups result from relative deficiencies in political and economic resources and differences in motives that determine whether workers have long- or short-term employment interests. Due to uneven development, migration from a rural-agrarian area to an urban-industrial area often provides the impetus for split labor markets.<sup>3</sup>

Minority migration may drive down prevailing wage scales or threaten labor organizations to the extent that employers hire workers from this cheaper and resource-weak labor pool.<sup>4</sup> As a response, dominant labor may attempt to exclude minorities by limiting direct competition. Dominant labor can limit competition by developing a caste system, which confines lower-cost minority labor to low-paying, low-status jobs. Another option is comprehensive trade unionism, which would effectively equalize wages and provide collective working-class benefits. However, the solidarity strategy requires that both groups give primacy to long-term class interests, a situation that is less likely to prevail in a split labor market (Bonacich 1972, pp. 555–58; 1975, p. 607; Marks 1981, p. 294).

Minority workers may have little interest in solidarity regardless of the union's strategy. If they are recent migrants, they often have fewer options or support networks, less information about local strikes or conditions, and a greater need for immediate wages. Additionally, racial networks and obligations may clash with solidarity, and racist ideologies skew and distort information (Boswell 1986). Employers may also try to divide dominant workers from minority workers by promising racial privileges (Peters 1994).

If minority workers are sojourners, migrants who intend to soon return to their place of origin (Siu 1952), then only short-term benefits are of interest. Strikebreaking becomes more attractive (immediate and higher

<sup>3</sup> Migrant minority workers may lack the skills appropriate to jobs at their new location, but competition is still intense if minority labor costs are low enough to compensate for lost productivity and still undercut dominant wages.

<sup>4</sup> Lindbeck (1993, p. 37) points out that experienced, incumbent workers' (insiders') jobs may be protected from outsider undercutting by labor turnover costs. This is especially true where insiders threaten to disrupt the production process by not cooperating with cheaper labor in the workplace, thus raising the cost to employers of hiring outsider labor (Lindbeck 1993, p. 38). In a strike situation, however, employers' first priority is to maintain production levels; labor turnover costs are far outweighed by the costs of halting production. Thus hiring outsiders to fill vacant positions becomes more rational.

income) and striking more costly (lack of resources and greater vulnerability to repression). Sojourners have a temporary orientation to returns from employment, such that they gain no long-term benefits from union organization and strikes (or long-term investments in human capital). Given a convenient exit, they will endure harsh working conditions and be less interested in collective voice. Their intent is to quickly save (or remit) a specific sum in order to purchase land, start a business, get married, or otherwise become established in their lower-cost homeland (Siu 1952; Bonacich 1972, p. 551). If wages rise, then their numbers may diminish, because they will more quickly achieve their objective (Berg 1961; Boswell and Jorjani 1988). They tend to be single young males who are least enmeshed in local networks and most mobile (Bonacich 1972, p. 550; Bonacich and Cheng 1984, p. 27).

Despite intentions, sojourners often fail to return to their homeland for a variety of reasons but continue to split the labor market until their temporary orientation ends. If they stay in one region, their tenure at a specific job may be short-term; even if they stay at a specific job, they have the option of returning. In short, sojourners' orientation to a job or locale is only economic and lacks the same degree of commitment that holds for long-term residents. As such, sojourners are the quintessential source of split labor markets. However, split labor market research has been overly focused on sojourning migration, which limits the historical applicability of the theory. We argue that all recent migrants (and especially single young males) share some of the characteristics of sojourners to the extent that they are poor but mobile.

### The Crucial Test

As a "crucial test," the analysis of post-World War I race relations by Bonacich (1976) and Wilson (1978) resulted in not only a test of split labor market theory but also a set of inductive conclusions about what factors to add or emphasize. While appropriate for understanding the original case, these factors have to be examined separately to determine their general validity. Four of the results merit consideration.

The first is the *primacy of sojourning migration*. As strikebreaking was directly connected to the migration of southern blacks, Bonacich (1972, 1975, 1976), Wilson (1978), and others failed to consider whether black strikebreaking took place without sojourning. Further, past research has not attempted to measure recent migration, which varied by city.

A second key factor is the *history of the dominant labor movement*. The black migration was so large and abrupt that the past history of race relations in the North was of overall minor consequence in the

massively changed context. Explaining abrupt changes is especially useful for theory testing in historical analysis. However, this abrupt change in race relations does not apply to labor relations. Sewell (in press) points out that sociological studies of history examine the effects of structure and conjuncture at the expense of explaining the reciprocal effects of events, especially how the particular sequence of events constrains the possibilities of subsequent occurrences. Of central importance for solidarity is the past history of union activity. A history of union successes or failures affects how people estimate risks and probabilities and also affects the resources of the union. These resources include physical resources, such as strike funds, and cultural resources, such as reputation, obligation, and habitual patterns of behavior. A long-standing consistent sequence of events is likely to produce normative rather than calculative responses (initially). Union solidarity may depend heavily on past success; risk aversion may inhibit solidarity where weak unions have a history of failed strikes.

Third, there is a need to focus on *state repression*. Burawoy (1981) contends that neglecting the state is a result of analyzing racial conflict only in terms of the labor market, which limits analyses to the interests of individuals or groups. Nonmarket contextual factors tend to be ignored, the most important being the class articulation of the state. During this period, the major variation in racial politics was between North and South, which obscured important local differences within the North. Yet a repressive local government increases the costs of labor organizing. Thus, patterns of strikebreaking might result more from state repression than the number of recent migrants in the labor market, while a comparatively pro-union local government might make solidarity more viable.

Finally, the fourth issue is an unstated *assumption that the absence of a split labor market explains interracial solidarity*. Historical analyses have pointed to interracial strikebreaking as the key cause of the steel strike's failure. However, particular cases of interracial solidarity were never studied explicitly in industrywide analyses, precluding a comparative analysis of strikebreaking versus solidarity. The theory has successfully been applied to similar cases, notably sojourning Asian immigrants (Bonacich and Cheng 1984; Boswell 1986), but while these studies place greater emphasis on the state and ideology, the centrality of sojourning reinforced the original focus. None of the split labor market studies concentrate on solidarity, and rarely have cases been compared (Boswell and Jorjani 1988).

If we are to explain when split labor markets produce strikebreaking and when workers can overcome the split through interracial cohesion, we need to expand the theory to include analyses of solidarity. The theory

should take into account the union history and the political context, explaining not just that they matter but when they matter most.

### Costs, Benefits, and Risks

Strikebreaking versus union solidarity provides a classic dilemma in studies of collective action (Olson 1965). The choice depends on workers' estimates of the costs, benefits, and risks, along with the probability of success and the possibility of free riding. Strikebreaking provides immediate employment at the cost of interpersonal and union sanctioning (possibly including violence). Strikebreakers risk job loss if the strike succeeds and lower wages, worse working conditions, and enhanced management control if the strike fails.<sup>5</sup> Solidarity, conversely, provides a higher benefit if the strike is successful and if the costs of striking can be recouped. However, strikers risk job loss and other costs if the strike fails.

Assessments of costs, benefits, and risks are more numerous and ambiguous than most participants can fully calculate. But even if a worker decides that a strike is warranted, the actions of a single individual have little or no effect. Because unions pursue collective goods, the possibility of free riding may provide little motivation for any individual to strike. In this sense, strikebreaking versus solidarity resembles a Prisoner's Dilemma, where individual behavior may be at odds with collective rationality.

However, individuals confront a union organization that has already decided to strike. Unions respond to the collective action problem by monitoring and sanctioning individual and group compliance and otherwise ensuring the solidarity of collective obligation (Hechter 1987; Taylor 1987, p. 9). Unions have the time and resources to calculate costs and benefits, even if individuals do not (Stinchcombe 1990). Union calculations may still be wrong, but decisions should predictably correspond to structural conditions. Success depends on a host of factors, such as union solidarity, company solidarity, resources, market conditions, state support or repression, and crosscutting (nonclass) racial and other interests.

While a number of other factors could be discussed, the condition we are concerned with is a racially split labor market. Workers are heteroge-

<sup>5</sup> McAdam (1986) points out that individuals can estimate and control costs more than risks, which depend more on the actions of others and the environment. As the probable cost of inaction or failure can be extrapolated from current conditions, determining the costs tends to be more certain than the probability of success or the amount of potential gain. We then might assume that consideration of the cost of inaction predominates in decision making. This might explain why unions embark on strikes when they seem unlikely to win.

neous in terms of the costs, benefits, and risks that are correlated with race, as per the theory's postulates. It may be rational for white workers to strike at the same time that it is irrational for minorities to strike. This is especially true if large numbers of minorities are recent migrants, whose higher discount rates inhibit collective action. The "shadow of the future" can facilitate cooperation in Prisoner's Dilemma situations where the importance of the next move relative to the present move is sufficiently high (Axelrod 1984, pp. 12–16; 1986). As short-term workers, however, sojourners discount the utility of future interactions at a higher rate, reducing the probability of cooperation.

#### A FORMAL MODEL OF COLLECTIVE ACTION<sup>6</sup>

Differences in costs, benefits, and risks create divergent worker strategies that reflect existing racial splits in the labor market. Actors' strategies are affected not only by short-term or long-term employment goals but also by the sanctions (rewards and punishments) the union organization offers and by the salience of intragroup norms that regulate members' behavior. The collective action problem in a split labor market can be modeled as an iterated mixed sanction system.<sup>7</sup> Such a model is appropriate because interactions are dynamic and sequential, because sanctions are directed at both the individual and the larger group, and because the working class cannot be treated as a corporate actor.<sup>8</sup> We are interested in the conditions that facilitate interracial union solidarity, so we do not focus on causes of union victory or defeat.

Following Heckathorn (1988, 1989, 1990a, 1990b), we model the collective action problem in terms of an *agent*, the union, that monitors and sanctions group members for defection and a heterogeneous group of *actors*, dominant and minority workers.<sup>9</sup> This is a mixed sanction system

<sup>6</sup> A supplementary paper (Brown and Boswell 1994) that provides a discussion of Heckathorn's formal theory and an explanation of our model, equations, modifications, and computational procedures is available from the authors.

<sup>7</sup> The system is iterated in the sense that actors sequentially select strategies after calculating opportunities, risks, and payoffs. In the model, the term "iteration" describes the strategy choice of a single actor. This definition differs from repeated game-theoretic models, where the game itself and all actors' moves are repeated at each iteration (e.g., Axelrod 1984, 1986).

<sup>8</sup> Bianco, Ordeshook, and Tsebelis (1990) reveal some of the problems involved in making predictions about repeated or *n*-person interactions based on one-shot, two-actor game-theoretic models.

<sup>9</sup> Ethnic (not just racial) heterogeneity may also foster divergence in worker strategies. While we acknowledge the diversity of working-class ethnic groups, the major split in the labor market was racial in the 1919 steel strike, and intraclass conflict occurred primarily between blacks and whites. Thus, our analysis focuses on racial, and not ethnic, dynamics. Subsequent research should address the impact of European immigrants on the development of ethnically split labor markets.

in that group members are subject both to *individualized sanctions*, rewards and punishments directed at individual actors, and to *collective sanctions*, which apply to all members of the actor's group (Heckathorn 1990a, pp. 366–67). Thus, individual actions have spillover effects that affect other group members (Heckathorn 1990a, p. 367). A single strike-breaker may be subject to individual sanctions, such as expulsion from the union, but also subjects other workers to a collective sanction by diminishing the effectiveness of strike activity or prompting an increase in the number of pickets. Spillover creates incentives for all workers to regulate one another's actions via *compliance norms*, which buttress the union's dictates (Heckathorn 1990a, p. 367).

Where individual sanctions are augmented by collective sanctions, the free-rider problem associated with the Prisoner's Dilemma may assume first- and second-order components (Heckathorn 1988, 1990a). The first-order free-rider problem involves the temptation to benefit from collective action without bearing the costs of providing public goods. In our case, strikebreakers may ultimately receive the higher wages and benefits associated with union victory without participating in the strike. The second-order free-rider problem refers to the temptation to benefit from intragroup compliance norms that curtail first-order free riding without sharing the costs of creating and maintaining these norms (Oliver 1980; Heckathorn 1990a, p. 368). Thus, a worker may free ride at the normative level by failing to discourage others from crossing the picket line.

Heckathorn (1989, pp. 81–82; 1990a, p. 368) has shown that in collective sanction systems actors are confronted with (at least) four choices. *Full cooperation* involves cooperation with the agent at the first level and support of intragroup compliance norms at the second; that is, refusing to cross the picket line and encouraging others not to cross. This strategy minimizes the probability that the actor or the actor's group will be subjected to external sanctions, but the actor must bear the individual costs of first- and second-level cooperation. *Hypocritical cooperation* is defection at the first level (strikebreaking) but participation in the intragroup sanctioning system (urging others to support the strike). Here, the actor free rides at the first level but cooperates at the second. *Private cooperation* involves cooperation with the agent at the first level but defection at the second. Finally, *full defection* is failure to cooperate with both the agent and the intragroup sanctioning system.

Additional alternatives become apparent when considering third-level behavior. Heckathorn (1990a, p. 377) notes that when intragroup compliance norms promote "compliance even when the individual costs exceed the collective gain" oppositional control may develop, where actors oppose the attempts of other actors to regulate first-level behavior. Thus, second-level compliance norms target first-level behavior, while third-

TABLE 1  
COOPERATION AT LEVELS 1-3 AND STRATEGIC CHOICES

LEVEL				STRATEGY	DEFINITION
1	2	3			
C	C	C	.....	(Self-defeating)	Illogical combination
C	C	D	.	Full cooperation	Strike* and encourage others to join (picket)
C	D	C	....	Hypocritical opposition	Strike but encourage others to strikebreak
C	D	D	...	Private cooperation	Strike and neither encourage nor discourage others
D	C	C	....	(Self-defeating)	Illogical combination
D	C	D	.....	Hypocritical cooperation	Strikebreak but encourage others to strike
D	D	C	.....	Full opposition	Strikebreak and encourage others to strikebreak
D	D	D	.	Full defection	Strikebreak and neither encourage nor discourage others

NOTE —C = cooperate, D = defect

\* We consider any worker who does not cross the picket line to be on strike, regardless of actual union membership.

level oppositional control represents a "meta-norm" (Axelrod 1986) aimed at affecting second-level behavior (Heckathorn 1990a, p. 378). For example, if membership dues (or other costs) greatly exceeded the benefits provided by the union, it would become rational for workers to oppose the (second-level compliance) norm of joining the union. Two of the four additional strategic choices introduced are theoretically interesting. The first, *hypocritical opposition*, refers to the strategy of cooperating at the first level and opposing compliance norms (third-level opposition). The second, *full opposition*, describes the strategy of defecting at the first level and opposing compliance norms through third-level opposition (Heckathorn 1990a, pp. 371-81).

All eight strategic choices are listed in table 1. Two are contradictory (both exercising and opposing compliance norms) and can be ignored. The remaining six strategies make up the set of options available to each actor. Actors are assumed to pursue the strategy that provides the maximum payoff. An actor's selection of a strategy (or "move") corresponds to a single iteration (or "period") of the model and the system variables change in accordance with each actor's choice. Thus, the structure of the system may differ at each iteration, and the optimal strategy for each actor is partially a function of the effects of prior selections of the other actors (i.e., each worker is affected by whether previous workers supported striking or strikebreaking). This model is sequential such that, where  $N$  equals the total number of actors, actor 1 makes the first move

at period 1, followed by actor 2 at period 2, until period  $N + 1$ , where actor 1 makes a second move (Heckathorn 1990b, p. 6). Ultimately, the system reaches equilibrium, where no actor's position can be improved through a unilateral strategy change.<sup>10</sup>

In Heckathorn's (1990a, 1990b) model, strategy selections are a function of the following fixed system parameters: the agent's ability to monitor defections ( $M$ ), the strength of individual sanctions that apply to individual defectors ( $S_i$ ), the strength of collective sanctions that apply to the group as a whole ( $S_c$ ), the efficacy of each actor's control over others at levels 2 and 3 ( $E_2$  and  $E_3$ ), and the costs to each actor of cooperating at each level ( $K_1$ ,  $K_2$ ,  $K_3$ ). Given these parameters, opportunities to cooperate and defect at each level and the risks and payoffs associated with each of the six strategies can be calculated (see Heckathorn 1990a, 1990b; Brown and Boswell 1994). Each actor's decision to cooperate or defect at levels 1, 2, and 3 will affect the opportunities, risks, and payoffs (system variables) facing subsequent actors and will thus affect their strategy selections.<sup>11</sup>

### Racial Groups and Heterogeneity

In recent work, Heckathorn (1991, 1993) explores the possibility that a group of actors may be heterogeneous with respect to values of the system parameters. We model overall group heterogeneity as a function of racial subgroup membership. Actors may be homogeneous within subgroups (the system parameters are the same for all actors in a subgroup) but heterogeneous between subgroups (the system parameters are different across subgroups). The number of subgroups ( $n$ ) is less than or equal to  $N$ . Where  $n = N$ , the system is perfectly heterogeneous and each actor has a unique set of parameters.

Split labor market theory argues that the two crucial labor market subgroups are the dominant (white) workers and minority (black) workers. For modeling purposes, we assume a labor force that is 70% domi-

<sup>10</sup> This result is similar to a noncooperative Nash equilibrium (Nash 1951; also see Heckathorn 1990a, p. 374) as discussed in game theory literature. Nash equilibria result when rational actors (1) select utility-maximizing strategies and (2) do so based on the anticipated strategy selections of all other rational actors in the system (Binmore and Dasgupta 1986, pp. 1–5; Kreps 1990, pp. 28–36). Thus, game theory is typically associated with "strategic interaction." While the equilibrium levels identified in our analyses result strictly from the choices of expected utility maximizers, the formal theory simulates strategic interaction in the sense that each actor's decision occurs in a context influenced by the prior decisions of all other actors.

<sup>11</sup> Derivation of the equations that relate the terms of the original model can be found in Heckathorn (1990a). Our modified equations are presented in the appendix.



nant workers and 30% minority workers. Additionally, another important difference between minorities is between local and migrant status. Thus, our model further divides minority workers into two subgroups, for a total of three subgroups (dominant, local minority, and migrant minority workers). Where minority labor is predominantly local, local minorities represent 20% of the total and migrant minorities represent 10%; where minority labor is predominantly migrant, locals represent 10% of the total and migrants represent 20%.

Heckathorn's formal model can be further adapted to the postulates of split labor market theory. Because the working class is heterogeneous with respect to race and recent migrant status, the costs of compliance ( $K_1$ ,  $K_2$ ,  $K_3$ ) and the efficacy of control at levels 2 and 3 ( $E_2$ ,  $E_3$ ) differ across the subgroups. Specifically, minorities (both local and migrant) are modeled as being subject to 5% greater collective sanctions and as having 5% higher costs of first-level compliance compared to dominant local workers. Also, compared to local workers, recent migrants have 20% higher compliance costs and a 20% lower efficacy of control at levels 2 and 3 owing to their need for immediate wages and their temporary status in the labor market (in addition to the 5% changes attributed to minority status).

Higher compliance costs for migrant minorities indirectly reflect the higher discount rate placed on the value of future outcomes. That is, the cost of migrant minority compliance is relatively higher due to the primacy of their immediate economic interests; they are less concerned about future payoffs and are therefore more likely to engage in first-level defection when it will yield an immediate (though possibly short-term) gain. These differences explain the tendency for racial conflict to undermine collective action, especially when industrial expansion promotes the migration of racially distinct groups.

Levels of solidarity (compliance) are affected by characteristics of the union as well as by characteristics of the workers. A strong union will have the resources to employ high levels of sanctions (both collective and individual) to deter defections. Because union strength is largely a function of the resources that accrue through the cooperation of its members, we modified Heckathorn's original formal theory to reflect a diminishing ability of the agent to monitor defections and apply collective and individual sanctions over time if the overall compliance level fails to cross a specified threshold. We posit an interaction between compliance levels and the ability of the external agent to enforce its dictates beginning with period  $N + 1$ . Thus, system parameters  $S_c$ ,  $S_i$ , and  $M$  diminish by 3% for each iteration beginning with actor 1's second strategy selection, if the overall compliance level is not greater than 50%. This has a dramatic rebound effect on levels of compliance over time and reflects the fact

that the initial success of union organizing may not be sufficient to ensure long-term cooperation. Because stronger sanctions increase the risks of and lower the payoffs for defection, strong unions will achieve higher compliance levels. In addition, strong unions will have a greater monitoring capacity; thus defectors will be more readily detected and punished. Parameter values for all groups can be derived from the information in table 2.

Congruent with our discussion of split labor market theory, we posit that mixed sanction systems with a strong union and a predominantly local minority labor group will produce the highest levels of compliance. Systems with a strong union and predominantly migrant minority labor or weak unions and predominantly local minority labor will produce intermediate levels of compliance. Finally, systems where a weak union is combined with predominantly migrant minority labor will produce the lowest levels of compliance. Figure 1 presents 40 iterations of the model using four separate sets of system parameters. Compliance levels are consistent with predictions derived from split labor market theory and reflect our theoretical expectations given variations in union strength and the percentage of minority migrants in the labor market.<sup>12</sup>

For the strong union/local minorities interaction, the group compliance level reaches .58 at equilibrium (see Heckathorn 1990a, p. 374). The remaining groups fail to exceed the .50 compliance level by the eleventh period, and union resources consequently begin to decline. As  $S_u$ ,  $S_m$ , and  $M$  fall, first-level defections increase and compliance norms begin to erode. Thus it becomes increasingly rational to oppose the normative system that supports first-level compliance. Once an actor has selected full opposition, compliance rapidly falls to zero and the exercise of full opposition becomes universal.

Our model strongly suggests that union strength plays a significant role in the development of worker solidarity in a racially split labor market. Solidarity is especially important over the long term for an industrial union, which organizes low-skilled workers throughout a given industry.<sup>13</sup> Over time, discrimination crowds minority workers into low-paying

<sup>12</sup> These results are quite robust; the pattern of high, medium, and low compliance levels depicted in fig. 1 obtains for a wide range of parameter values. All base parameters can be increased or decreased by at least 25% and the predicted pattern of compliance levels holds (although the specific compliance levels are affected by fairly modest parameter changes). A more elaborate discussion of both the model's robustness and measures of parameter sensitivity is undertaken in our supplemental paper (Brown and Boswell 1994).

<sup>13</sup> Unlike industrial unions, craft labor, especially when trained through apprenticeship, may be able to maintain discrimination based on monopoly over scarce skills. Craft labor may have less to gain from solidarity and less to lose over the long run (although a skill monopoly does encourage deskilling).

TABLE 2  
PARAMETERS FOR DOMINANT AND MINORITY SUBGROUPS ( $N_x$ ,  $N_y$ ,  $N_i$ ) EXPRESSED AS A PROPORTION OF BASE-LEVEL PARAMETERS

	$S_c$	$S_i$	$M$	$E_2$	$E_3$	$K_1$	$K_2$	$K_3$
Base parameters .....	100	925	.1	.2	.2	100	6	6
Strong union:								
$N_x$ (dominant) .....	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$N_y$ (local minority) ..	1.05	1.00	1.00	1.00	1.00	1.05	1.00	1.00
$N_i$ (migrant minority) .....	1.05	1.00	1.00	.80	.80	1.25	1.20	1.20
Weak union:								
$N_x$ (dominant) .....	.75	.75	.75	1.00	1.00	1.00	1.00	1.00
$N_y$ (local minority) ..	80	75	.75	1.00	1.00	1.05	1.00	1.00
$N_i$ (migrant minority) .....	80	.75	.75	80	80	1.25	1.20	1.20

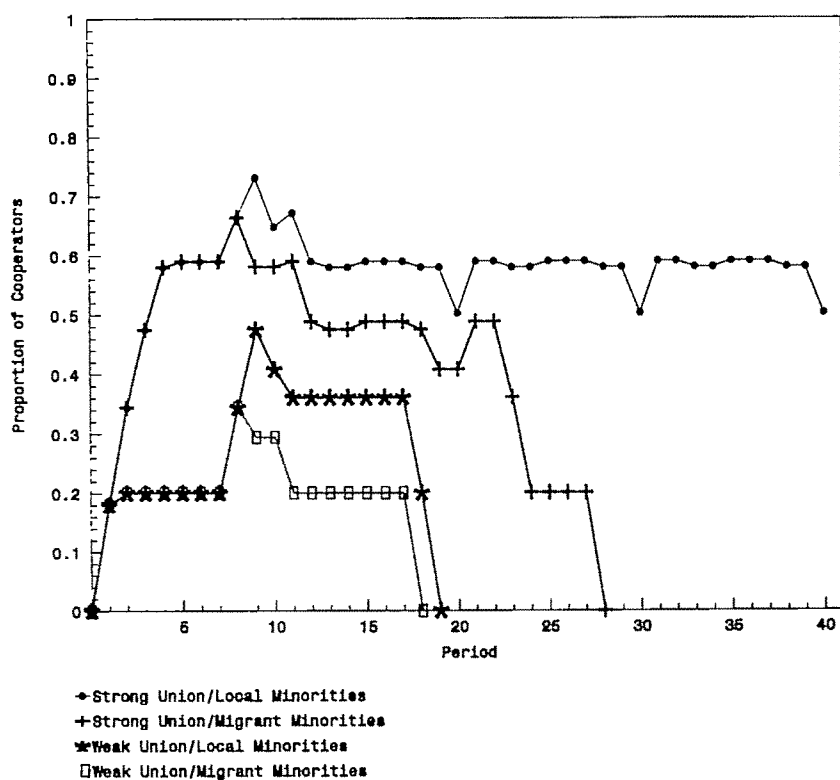


FIG. 1.—Formal model of labor market group compliance levels

jobs, reproducing the split and creating further incentives for employers to hire strikebreakers (Boswell 1986). Discrimination results in lower relative wages for dominant labor (as well as minority workers) by retarding unionization, dividing working-class politics, and disrupting investments in human capital (Reich 1981). On the other hand, if industrial unions pursue solidarity (and the labor market remains stable), then competitive pressure should eventually erode the split. Based on our formal theory of collective action, we expand split labor market theory to predict higher levels of solidarity where recent migrants are few and where the union is comparatively strong.

We subject the expanded theory to a rigorous test over three issues. First, we determine if the assumptions and inductive emphases derived from past research apply to the community level of analysis, where the events actually took place and where distinct comparisons can be made. Second, we test the importance of the recent migrant thesis from split labor market theory against competing labor history and state repression

alternatives. Third, we determine whether the theory applies to solidarity as well as strikebreaking according to expectations derived from our formal analysis. The strike as a whole was unsuccessful for many reasons, black strikebreaking being the crucial factor. But the extent of strikebreaking differed substantially from one community to the next, and in two cases interracial solidarity prevailed for the duration of the strike.

## HISTORICAL OVERVIEW

The period between 1910 and 1920 witnessed an unprecedented northward movement of southern blacks. The push/pull factors that initiated and sustained this migration have been well documented and include shifts in the demand for labor, the wartime industrial surge, agricultural problems in the South, regional wage differentials, restricted European immigration, and southern racial oppression (see Scott 1920; Donald 1921; Lewis 1931; Myrdal 1944; Henri 1975; Fligstein 1981; Gottlieb 1987; Grossman 1989; Lemann 1991). Four hundred thousand blacks, or 5% of the total minority population, relocated from the South to the North between 1916 and 1918 (Marks 1983, p. 73; Scott 1920, p. 3). In addition, the number of black farm workers declined by more than 24% during this period (Dickerson 1986, p. 29). Black migration during World War I was largely a male phenomena (Donald 1921, p. 408; Florant 1942, p. 747; Marks 1983, p. 76). Because most of these migrants sought higher-paying urban jobs, the migration represented a shift from agricultural sharecropping to the ranks of the industrial proletariat.

By 1920 black workers in iron and steel numbered over 125,000 (Spero and Harris [1931] 1969, pp. 84–85). However, blacks were predominantly assigned to low-paying, unskilled positions and often performed the most dangerous and strenuous tasks.<sup>14</sup> Table 3 shows the number and percentage of blacks and whites born out of state but residing in various northern steel-producing cities in 1920. Comparing the two groups suggests a much higher level of mobility among blacks.<sup>15</sup> Further, in all the cities the number of blacks born out of state is at least twice

<sup>14</sup> For information on working conditions, see Jones (1919), "Southern Negro" (1924, pp. 43–44), Spero and Harris (1969, pp. 153–54), and Dickerson (1986, pp. 49–52), who notes that often, when black steelworkers were employed in semiskilled positions, "the work and pay differed very little from unskilled jobs" (p. 264, n.81).

<sup>15</sup> Many authors have noted the transitory nature of black employment and high job turnover rates for black workers during this period, particularly among recent migrants to the North (see Spero and Harris 1969, pp. 178–79; Piore 1979, pp. 158–59; Gottlieb 1978, pp. 206–8; Grossman 1989, pp. 195–97). We also note that table 3 reveals variation in out-of-state births for whites by city that may reflect differences in levels of European immigration.

TABLE 3

NUMBER AND PERCENTAGE OF POPULATION BORN OUT OF STATE FOR SELECTED  
NORTHERN STEEL-PRODUCING CITIES, 1920

CITY	BLACES		WHITES	
	N	%	N	%
Akron, Ohio .....	4,736	85.4	76,915	46.6
Chicago, Illinois* .....	91,095	84.0	401,630	22.5
Cleveland, Ohio* .....	27,907	81.9	112,804	21.6
Columbus, Ohio .....	13,484	60.9	32,542	16.4
Dayton, Ohio .....	5,998	66.7	26,657	20.4
Gary, Indiana* .....	4,924	93.5	18,375	54.7
Pittsburgh, Pennsylvania* ....	26,106	69.7	47,764	11.1
Toledo, Ohio .....	4,106	73.4	50,237	25.2
Youngstown, Ohio* ..	5,488	82.8	29,564	32.2

SOURCE —U.S. Bureau of the Census (1922, pp. 661-65).

\* Involved in 1919 steel strike.

the number born in the state, suggesting that the majority of the black population had migrated, probably from the South.

The racial composition of unions began to reflect the movement of blacks to the cities. Total black membership in American trade unions was slightly over 30,000 in 1900 and close to 70,000 (about 3% of all workers) in 1910 (Spero and Harris 1969, pp. 76-78). In the steel industry, however, unskilled workers and blacks were generally excluded from the unions.<sup>16</sup> The largest of the steel industry unions was the Amalgamated Association of Iron, Steel, and Tin Workers. In 1891, prior to the disastrous Homestead strike, the Amalgamated Association had a national membership of 24,000 (Newell 1961, p. 122). However, a series of labor defeats (particularly in 1892, 1901, and 1909-10) had virtually driven it out of the industry by 1910. Once the Amalgamated Association (along with smaller steel unions) had been defeated, the U.S. Steel Corporation had extensive control over hours, wages, and working conditions (Brody 1965, p. 27; Yellen 1974, p. 252).

Although blacks were generally excluded from unions, dominant workers encouraged blacks to go on strike, while employers recruited blacks as strikebreakers.<sup>17</sup> Whether blacks joined unions or crossed the picket

<sup>16</sup> See Dickerson (1986, pp. 7-27) for an overview of the early history of black workers and steel unions.

<sup>17</sup> See Bonacich (1976, p. 41) for a list of strikes between 1916 and 1934 where black strikebreakers were used. Foster (1920, p. 207-8) reports that during the 1919 steel strike, blacks were often taken to strike areas under false pretenses (also see Gottlieb 1987, p. 57).

line, underlying wage disparities and inequality remained. In addition, some black leaders and organizations (such as the Pittsburgh Urban League) advocated strikebreaking ("The Black Man and the Unions" 1918; Spero and Harris 1969, pp. 128, 139–42, 258–59).

Between 1914 and 1919, strikes in the steel industry increased in size and number. Strikes in iron and steel climbed from 14 in 1914 to 76 in 1919, and the number of workers involved in all strike activity exceeded 4 million in 1919, more than twice the number involved in strikes for any year between 1916 and 1926 (Whitney 1920, p. 200; Peterson 1937, pp. 38–39). The industrial boom and labor shortage generated by World War I gave steelworkers new leverage, and the American Federation of Labor (AFL) began efforts to organize the industry in 1918. The AFL voted to permit the entrance of black workers in June of 1919, and the National Committee for Organizing Iron and Steel Workers set a precedent by making an effort to incorporate black workers into the developing steel union. However, blacks were segregated or excluded from some of the local labor organizations, despite efforts at the national level to ameliorate racial divisions (Brody 1965, p. 162).

After approximately 18 months of organizing efforts and in response to a flurry of localized labor disturbances that threatened to fragment the developing national organization, the national committee called for a union vote on a proposed work stoppage. On August 20, 1919, 98% of the workers voted in favor of a strike in the event that their demands were not met (Foster 1920, p. 78). The chief objectives included securing the right to unionize, a reduction in the length of the work day and work week, and wage increases (Gadsby 1919; Interchurch World Movement 1920, p. 246; Brody 1965, pp. 100–101). Because the steel industry (led by U.S. Steel)<sup>18</sup> refused to acknowledge the legitimacy of the union or grant concessions, the national committee called for a strike on September 22. Despite President Wilson's efforts to mediate the dispute, the conflict was not resolved and the strike commenced.

William Foster (1920, p. 100), secretary-treasurer of the national committee, reported that over 365,000 workers (about half of the total) had left their jobs one week into the strike, and in many plants production was completely halted. Labor organizers had hoped that sympathy strikes among other industrial workers would follow, but such movements failed

<sup>18</sup> U.S. Steel developed in 1901 and involved a merger of previously independent steel producing companies at the national level. According to Foner (1973, p. 12), U.S. Steel was the first billion-dollar corporation and combined 11 massive steel, mining, and shipping companies that controlled 60% of the steel market and many of the national industries involved in metals, fuel, and transport (also see Garraty 1960; Edwards 1979, p. 44).

to take hold on a large scale. Despite initial success, the movement began to crumble and workers gradually returned to the mills. In December, 109,300 workers were still on strike, and the production of steel had risen to between 50% and 60% of normal (Brody 1965, p. 174). Although the strike remained fairly strong in a few areas (such as Cleveland, Wheeling, and Buffalo), it had become apparent by the end of December that a strike on a national scale could not continue. On January 8, 1920, the strike was formally called off and Foster resigned.

National-level factors contributed to the failure of the strike and the union movement. The labor organization as a whole became fragmented, because skilled workers were particularly reluctant to strike (Foster 1920, p. 249; Interchurch World Movement 1920, p. 248; Asher 1978, p. 61). The financial strength of the steel companies was substantial (due to inflated wartime production), enabling owners to better weather the strike (Asher 1978, p. 82). Rumors of a Bolshevik conspiracy proliferated, diverting attention from the grievances of the workers and turning public opinion against a strike thought to be incited by foreign radicals (Foner 1988, p. 164; Murray 1951, 1955).

However, local factors had the greatest influence on the outcome of the strike in particular plants. The steel industry was able to undermine the efforts of the workers through violent intimidation and the widespread use of strikebreakers (Foster 1920, pp. 207–8; Interchurch World Movement 1920, p. 248; Murray 1951). Communities with a history of failed strikes also tended to exhibit lower worker solidarity than other communities (Foster 1920, pp. 103, 181; Asher 1978). Additionally, the political influence of steel officials resulted in the mobilization of local and state authorities on behalf of the industry in key steel towns (Murray 1951, p. 456).

The effectiveness of strikebreakers, especially blacks imported from the South and numbering as high as 40,000, is frequently cited as the crucial cause of the union's defeat (see Foster 1920, p. 207; Yellen 1974, p. 288; Foner 1974, p. 144; Dickerson 1986, p. 92). The importation of black labor during the strike varied substantially by city and was most prominent in breaking the strike in the Pittsburgh area (Spero and Harris 1969, pp. 261–62; Dickerson 1986, pp. 88–93).

The communities that were involved in the 1919 strike differed substantially in terms of the key factors hypothesized to affect biracial labor solidarity. In only two cases, Cleveland and Wheeling, did blacks and whites form a cohesive strike force. Black strikebreaking was prevalent in a number of locations including Buffalo, the Chicago area, Gary, Johnstown, Pittsburgh, and Youngstown. A high level of anti-union sentiment in the local government was apparent in most of the communities of the Pittsburgh area, as well as in other places, such as Bethlehem,



Buffalo, and Gary. Failed strikes (indicating weak union locals) had occurred in Bethlehem, Joliet, McKeesport, Pittsburgh, and other areas. The following section describes the fundamentals of QCA, which we use to discern the causes of strikebreaking and solidarity across the cases.

#### A COMPARATIVE APPROACH

Comparative analysis is particularly suited to this study, because the cases are fraught with multiple overlapping factors that tend to confound related, but distinct, causal mechanisms. As Ragin (1987, pp. 25–27) points out, the advantage of case studies is in explaining outcomes that depend upon multiple influences operating together within a specific context, which underscores the conjunctural nature of causation. QCA (Ragin 1987) is a methodology that addresses issues of causal complexity and lends itself to the analysis of qualitative data with a relatively small number of cases.<sup>19</sup> This method is well suited for testing deductive expectations from our formal model.

In essence, QCA is simply an algorithmic application of the logic of comparative analysis (cf. Przeworski and Teune 1970 with Ragin 1987), which allows one to handle more complex relationships and a larger number of cases than is typically true in comparative historical research. More important, QCA forces the analyst to consider all possible combinations of causal factors (even those not empirically present), thus exposing hidden assumptions. This inhibits selection on the dependent variable or using ad hoc explanations of anomalies, problems endemic in case study research (Liebersson 1991).

QCA involves five key steps: (1) identify positive (or “present”) cases associated with a particular dependent variable and negative (or “absent”) cases associated with its absence, (2) identify and define the factors expected to influence outcomes of the dependent variable, (3) code variables for each case in a dichotomous manner of “presence” or “absence,” (4) construct a truth table that lists all possible combinations of variables and outcomes, and (5) use Boolean logic to minimize the truth table by eliminating redundant or superfluous cases (Ragin, Mayer, and Drass 1984, p. 228).

Our unit of analysis is the steel-producing communities where relevant variation in race relations occurred. Cities in the analysis have the following characteristics in common: the population was over 25,000, the steel industry was an important employer, black workers were a significant part of the steel workforce, and the recruitment efforts of the national

<sup>19</sup> For additional examples and debate concerning this approach, see Ragin (1991), Ragin et al. (1984), Liebersson (1991), and Boswell and Brown (1994).

committee reached each location during the organizing drive. We selected cases from Foster's (1920) account of the organizing drive and include only those locations that actually participated in the strike.<sup>20</sup> The final group of 16 cases ranges in size from single-industry towns to major multi-industry cities. We have completed a case study that describes strike events for each city, which space does not allow us to present here.<sup>21</sup> Efforts to organize the industry focused primarily on plants affiliated with U.S. Steel and, to a lesser extent, Bethlehem Steel. It was only in these two companies that workers responded heavily to union recruitment (Foster 1920, pp. 65–66, 100–101).

We derived five independent variables for the analysis that follows, one based upon the recent minority migrant hypothesis from split labor market theory, a second that focuses on union strength, a third that considers city size, a fourth that examines ownership of the steel companies, and finally, an alternative political power thesis. The dependent variable is divided into four categories (based on compliance levels from fig. 1): strikebreaking, union solidarity, and two mixed cases where neither strikebreaking nor solidarity dominates. First we focus on the main two variables that operate at the level of the split labor market, recent migration and union strength.

### Recent Migration

For recent migration to affect the course of the strike, a city must have had a relatively large population of young, single, black males that was increasing at a rate higher than that of the general population. The recent migrant variable, calculated for each case from the censuses of 1910 and 1920, is an index of three measures: black males 20–44 years old as a percentage of all such males (1920), the ratio of the rates of increase for black males compared to total males (1910–20), and the magnitude of the increase in black males (1910–20). Because marital status is also important (but not broken down by age), we calculated the same index for single black males.<sup>22</sup> The higher these measures, the higher the pro-

<sup>20</sup> A dearth of census and historical information prevents us from including cities with populations under 25,000 in our analysis. An advantage of this is that small, possibly idiosyncratic cases are excluded.

<sup>21</sup> An overview of strike developments by city and historical data pertaining to the coding of variables are available from the authors.

<sup>22</sup> Measures grouped by age reflect the age breakdowns in the census. Cities with larger populations give breakdowns in five-year increments, while cities with smaller populations collapse the age groupings. To maintain consistency across cities, we used the 20–44 year old category as a measure of relative "youth." We provide the data in tables B1 and B2 in the appendix. The percentage of the population born out of state is not available for many of the cities in the analysis, nor is this statistic reported

portion of recent black migrants with sojourning or temporary characteristics. While the characteristics "young" and "single" do not necessarily imply return migration, they are associated with transient or mobile workers. The historical record indicates that the initial wave of the black migration was disproportionately young and unattached labor; our analysis investigates the impact of these migrants on regional differences in interracial labor solidarity.<sup>23</sup>

We used the following criteria to evaluate each case: (1) whether a city's population of young males (20–44 years old) is more than 5% black in 1920; (2) whether the relative rate of increase is greater for blacks than for whites;<sup>24</sup> and (3) whether a city showed an increase of greater than 100% in its population of young black males between 1910 and 1920. These criteria ensure a clear diversity consistent with available historical literature. Cities that meet at least two of these three criteria for both the age and marital status components are coded "present" (given a value of "1") for the recent migrants with sojourning characteristics variable. The coding for each case is given in the last column of appendix tables B1 and B2. Note that the coding is the same across both age and marital status components of the measure, because most single males are also young.

The recent migrant variable, as defined, assumes a high correlation between the influx of recently arrived minorities into the steel industry and into the city as a whole, since the data are at the city level. A large proportion of the black migrants did obtain employment in the steel industry as the migration swelled. Census data show that the number of

---

by race or sex for 1910 or 1920. We acknowledge that the growth of the minority male population by city may be *caused by* (through labor recruitment) as well as a *cause of* black strikebreaking, but this issue is not central to our argument. That high concentrations of migrant minorities are coincident with minority strikebreaking supports our hypotheses in either case. Gottlieb (1987, p. 57) and others (Epstein 1918, p. 25; Henri 1975, p. 80), however, do suggest that labor recruitment played a relatively small role in the migration.

<sup>23</sup> The migration began abruptly with the commencement of the war, reached a peak between 1916 and 1918, then dropped off slightly in the years between 1919 and 1920 (see Scott 1920; Donald 1921; Kennedy 1930; Marks 1983). Available evidence does not suggest that this pattern deviated substantially by city.

<sup>24</sup> We calculated the relative rate of population increase as

$$\left[ \frac{(BMA^{1920}/TBM^{1920})/(TMA^{1920}/TM^{1920})}{(BMA^{1910}/TBM^{1910})/(TMA^{1910}/TM^{1910})} \right] - 1,$$

where *BMA* = "black males 20–44 years old," *TBM* = "total black males," *TMA* = "total males (all races) 20–44 years old," and *TM* = "total males (all races)." A positive ratio indicates that the number of blacks 20–44 years old increased at a higher rate than did the number of their white counterparts.

black males working in the steel industry of the urban North increased by 600% between 1910 and 1920.<sup>25</sup> Black migration to steel cities is statistically coincident with increases in the number of blacks employed by the steel industry. The percentage of black males 20–44 years old and the percentage of single black males correlate highly ( $r = .91$  and  $r = .93$ , respectively) with the percentage of black steelworkers for the 15 northern cities for which data are available in 1920. Using 1910 data for the same cities, these correlations are almost identical to those obtained using 1920 data (for both,  $r = .91$ ). Furthermore, the change (1910–20) in the number of black males 20–44 years old and the change in the number of single black males also correlate highly (for both,  $r = .89$ ) with the change in the number of black steelworkers over the 10-year period. The strength of these relationships suggests that increases of recent migrants in the urban labor market are consistent with increases in the population of black steelworkers.<sup>26</sup>

### Union Weakness

Union strength is reflected in the strike history in a given city. Although the 1919 strike was the first national-level attempt to organize industrial steelworkers, localized strikes occurred relatively frequently over the preceding 20 years. While workers occasionally won minor concessions, outcomes largely favored the owners (Foster 1920, pp. 8–15; Brody 1965). Past defeats indicate the relative strength of capital over labor and lower the expected probability for current success.<sup>27</sup>

We expect cities with a history of labor defeats to have a low propensity to generate biracial labor solidarity (or a working-class movement in general) due to “historical consciousness.” Asher’s (1978) interviews with participants in the 1919 strike in the Pittsburgh area suggest that the

<sup>25</sup> We based this calculation on census data for the 15 northern cities for which the number of black male steelworkers was available in both 1910 and 1920 (Bridgeport, Conn.; Buffalo, N.Y.; Chicago; Cleveland; Columbus, Ohio; Detroit; Milwaukee; New Haven, Conn.; New York City; Newark, N.J.; Philadelphia; Pittsburgh; Syracuse, N.Y.; Toledo, Ohio; Worcester, Mass.). This information is not reported for cities with populations below 100,000. Of these 15 cities, five (Buffalo, Chicago, Cleveland, Milwaukee, Pittsburgh) participated in the 1919 strike.

<sup>26</sup> In smaller industrial towns the steel industry was often one of the primary sources of employment, thus correlations between recent black migration and black entrance into the steel industry should be higher than for larger cities (where employment options are more diverse). While the measures contained in the recent migrant variable can be partially justified on theoretical and empirical grounds, the limitations described above should be considered when interpreting the results.

<sup>27</sup> A strong union need not strike so much as threaten to strike; past history makes the threat credible (Korpi and Shalev 1979).

lengthy history of defeats convinced many steelworkers that a national unionization effort would never be successful. Foster (1920, p. 181) notes that the ineffectiveness of the 1919 strike in the Bethlehem Steel plants was largely due to the "failure of previous strikes." General participation and commitment to the 1919 strike was weakened in cities where workers strongly questioned the possibility of success. A history of failed strikes may also point to low levels of organizational resources.

Because the 1919 steel strike was organized mainly in opposition to U.S. Steel, the history of failed strikes is coded "1" (present) if the steelworkers of a given city experienced a failed strike subsequent to the formation of U.S. Steel in 1901. Failed strikes that occurred before this time are less likely to have had a lasting impact on the consciousness of the local workforce, in part due to the attrition of many of the workers who participated in these strikes and in part due to the dramatic reorganization of capital represented by the formation of U.S. Steel. In cities with a history of failed steel strikes, we expect that dominant workers and labor organizers may less vigorously recruit minority workers due to the underlying sense of futility and their lack of power and resources relative to capital. Relatedly, if dominant workers are ambivalent about going on strike, then minority workers confronted with a history of poor relations with the labor movement should be even less likely to join the union and participate in a work stoppage.

#### Outcomes: Solidarity versus Strikebreaking

The 1919 strike as a whole was characterized by black strikebreaking, which occurred in many, although not all, locations. Many cities had low levels of participation by white workers as well as black, and the strike effort was weak from the beginning in these locales. Notable instances of biracial labor solidarity developed in only two cities, Cleveland and Wheeling. In short, low levels of minority compliance or active strikebreaking held in all but two cases.<sup>28</sup> Below we undertake a comparative analysis of the causal factors considered in the formal model presented in figure 1 (minority migration and union strength). We consider other factors only after assessing the most parsimonious model of split labor market precepts.

<sup>28</sup> Split labor market theory posits a strong relationship between racial antagonism and minority strikebreaking. We were unable to measure variation in preexisting evidence of racial antagonism in a reliable manner. We considered race riots, but found only one instance (Chicago) shortly preceding the strike. Race riots had occurred in other cities not included in this study, and we can assume a high level of racial suspicion and distrust in all cases.

TABLE 4

POSSIBLE COMBINATIONS OF INDEPENDENT VARIABLES AND DISTRIBUTION OF CASES

<i>M</i>	<i>U</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	Cases
1	1 . . .	1	0	0	0	East Chicago, Pittsburgh, Youngstown
1	0 . ....	1	0	0	1	Buffalo, Chicago, Cleveland, Gary, Johnstown
0	1 . .	0	0	1	0	Bethlehem, Joliet, McKeesport, Milwaukee, New Castle, Reading
0	0	0	0	1	1	Decatur, Wheeling

NOTE.—*M* = recent black migrants (1 = higher proportion), *U* = union weakness (1 = historically weak), *A* = whites support union/blacks strikebreak, *B* = whites strikebreak/blacks support union, *C* = biracial strikebreaking low overall compliance, and *D* = biracial labor coalition high overall compliance

## CASE COMPARISON

The task, then, is to determine what combination of factors promoted or inhibited strikebreaking versus class solidarity for the 16 cases in the analysis. The variables suggested by the formal model are presented in table 4. Following Ragin (1987), the presence of a variable is represented by a "1" in the table (and by a capital letter in the equations), and its absence by a "0" (and a lowercase letter).

Note that two contradictions prevent us from clearly determining the underlying causal combinations and from performing a reliable QCA. Buffalo, Chicago, Gary, Johnstown, and Cleveland all have the same combination of the independent variables (*Mu*). However, Cleveland differs from the other four in terms of the dependent variable, where workers achieved interracial solidarity. Second, Wheeling and Decatur have the same combination for both variables (*mu*), although only Wheeling workers developed a biracial labor coalition. If the contradictory cases (Cleveland and Wheeling) are ignored, all cities that experienced black strikebreaking are associated with higher proportions of recent minority migrants. While this supports the major precept of split labor market theory, the contradictions indicate that the model, as specified, inadequately accounts for outcomes across all cities.

Contradictions can be resolved inductively by returning to the original cases and looking for additional variables that directly address the contradictory results (Ragin 1987, p. 113). Though we consider union (strength/weakness) and minority (migrant/local) characteristics to be the primary factors that determine the possibility of cooperation at the level of the labor market, it is likely that this interaction is conditioned by contextual factors that our present model fails to take into account. Below we explore three such contextual variables—city size, ownership of the steel companies, and local government repression—and examine their impact on the QCA results.

### City Size

Larger cities might pose greater organizational problems for unions and might also be more likely to draw temporary migrants than smaller locations. Thus, we might resolve the anomalies in table 4 by including this variable in the analysis, in addition to recent migration and union strength. We operationalize this variable in a simple manner, coding those cities with populations greater than 100,000 in 1920 (Buffalo, Chicago, Cleveland, Milwaukee, Pittsburgh, and Reading) as "1" for *P* (large population) and all others as "0" (*p*).

However, in an analysis not shown, city size fails to elucidate the contradictions in the data. The resulting three-variable table still contains the same two contradictions, Cleveland and Wheeling, thus precluding the derivation of any reduced equations. We can generally assume that differences in city size are irrelevant.

### Company Ownership/Employer Status

Foster's (1920, pp. 181–84) account of the organizing drive indicates that strike participation tended to be much greater in cities with U.S. Steel plants than in cities dominated by Bethlehem Steel. At the same time, U.S. Steel's size gave it a high level of economic resources and political influence. This difference may explain some of the variation in interracial solidarity and resolve the contradictions in the data. Bethlehem and Reading, the only Bethlehem Steel cities in the analysis, are coded "1" for *I* (industry ownership) and all others as "0" (*i*) for U.S. Steel. However, contradictory combinations again emerged when *I* was incorporated into the model (not shown), precluding the derivation of equations.

### Local Government Repression

A final important context is the level of political repression that exists in a particular city. If the steel industry can use political power to thwart organizing efforts, it will retard the formation of labor coalitions by inhibiting the ability of unions to overcome the barriers needed to win black support and ameliorate existing racial divisions. In a split labor market situation, the repressive actions of anti-union local governments will further fragment the labor movement along racial lines. Dominant unions need greater resources and more reliable networks of communication to organize blacks than whites due to organized labor's history of poor race relations and the ubiquity of racist ideologies. Repressive local governments make organization of blacks even more problematic, raising the

risk of organizing. Fewer workers will join the union, and strikebreaking may be seen as more acceptable. The widespread use of violence to suppress the organizing drive in many cities and the political weakness of (and lack of protection for) blacks relative to whites may also impede efforts to achieve biracial labor solidarity where local government repression is high.

Given the historical accounts of labor organizing at the local level (Foster 1920; Interchurch World Movement 1920; Brody 1965), along with the red scares of the time and local economic dependence on steel production, we can generally assume local state hostility toward the union unless there is clear evidence to the contrary. Local governments have been classified as repressive where local government officials (or their agents) prohibited labor organizing meetings, infringed upon free speech, or engaged in the systematic jailing, intimidation, beating, or murder of labor organizers or union members.<sup>29</sup>

Pro-union governments, on the other hand, intervened on behalf of the strikers (to prevent company repression or the importation of minority strikebreakers) or had been elected on the basis of strong labor platforms (Brody 1965, p. 152; Boryczka and Cary 1982, p. 165; Foster 1920, p. 184). We know from the case study research that Cleveland, Wheeling, and Milwaukee had relatively pro-union municipal governments and that in these cities the repressive tactics characteristic of western Pennsylvania were notably absent. For example, the actions of the Cleveland mayor were decisive in preventing the entry of strikebreakers imported by the mill owners and in Wheeling the relatively pro-union stance of the local government had been effected by generations of strong union activity (Foster 1920, p. 184). The pro-union stance of Milwaukee's local government is suggested by Trotter (1985, p. 44), who notes that labor consistently supported the socialist mayor, who ran city government from 1916 to 1940. The source of pro-union governments is likely to be strong unions. Thus we can expect an interaction between the union and the state variables.

The revised truth table (table 5) shows that the previously contradictory cases are resolved by the inclusion of this variable. However, there is now one hypothetical case (*NUr*), that is not accounted for in the data (we address this below). With the contradiction resolved, we can now derive QCA equations for each of the outcomes.

The reduced equations are called "prime implicants," because they

<sup>29</sup> Initial results and case histories suggested that the key factor was whether the local government sided with labor during the strike. As a result, our coding treats neutrality or repression as "1" and support as "0" to isolate those governments that had a favorable stance toward labor.



TABLE 5

POSSIBLE COMBINATIONS OF INDEPENDENT VARIABLES AND DISTRIBUTION OF CASES

<i>M</i>	<i>U</i>	<i>R</i>		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	Cases
1	1	1	.....	1	0	0	0	East Chicago, Pittsburgh, Youngstown
1	1	0	.....	?	?	?	?	Hypothetical
1	0	1	.....	1	0	0	0	Buffalo, Chicago, Gary, Johnstown
1	0	0	.....	0	0	0	1	Cleveland
0	1	1	...	0	0	1	0	Bethlehem, Joliet, McKeesport, Milwaukee, New Castle, Reading
0	1	0	... ..	0	0	1	0	Milwaukee
0	0	1	....	0	0	1	0	Decatur
0	0	0	.....	0	0	0	1	Wheeling

NOTE.—*M* = recent black migrants (1 = higher proportion), *U* = union weakness (1 = historically weak), *R* = local government repression (1 = repressive or neutral), *A* = whites support union/blacks strikebreak, *B* = whites strikebreak/blacks support union, *C* = biracial strikebreaking: low overall compliance, and *D* = biracial labor coalition: high overall compliance. Resulting QCA equations are  $A = MR (+ MU)$ ,  $B = ?$  (no observed cases),  $C = mU + mR$ , and  $D = ur$

imply all of the causal combinations associated with the dependent variable (Ragin 1987, pp. 95–98). All of the cities that experienced high levels of black strikebreaking (*A*) had recent black migrants with sojourning characteristics and a nonsupportive local government. Not surprisingly, white strikebreaking with black union solidarity (*B*) did not occur over the length of the strike. In several cases, an initially high level of white solidarity was undermined by either weak unions or a repressive state where minorities were mainly local (*C*). Finally, biracial labor coalitions (*D*) developed where there was both a pro-union government (*r*) and a strong union (*u*). The results in table 5 support the recent black migrant hypothesis with respect to black strikebreaking but suggest that biracial labor solidarity may prevail despite an influx of recent black migrants in cities with favorable governments and strong unions.

While recent black migration was necessary for the development of black strikebreaking, its absence was not sufficient for explaining solidarity. Further, the contextual effects of a repressive government on black strikebreaking also appear to be necessary but only in combination with an incoming black migrant population. Union strength is important for the development of a biracial labor coalition. More interesting, perhaps, is the fact that the government context appears to play a critical role for determining every outcome.<sup>30</sup>

A problem with table 5 is limited diversity (Ragin 1987, pp. 104–13).

<sup>30</sup> Using DeMorgan's law (Ragin 1987, pp. 98–101), we can also derive equations for *a*, *b*, *c*, and *d* or for the absence of each outcome. The resulting equations are  $a = m + r$ ;  $b = ?$ ;  $c = M + ur$ ; and  $d = U + R$ .

One potential variable combination ( $MUr$ ) is unaccounted for in the historical data, signified by “?” in the outcome columns. By not including this hypothetical case in the analysis, we have implicitly assumed that it would not have produced an outcome (“0” in all equations). To be sure, some theoretical combinations may be absent in the data because they “combine incompatible elements and are therefore unlikely ever to exist” (Ragin 1987, p. 109). However, it would be imprudent to make this assumption based on only 16 cases. Certainly, municipal pro-labor governments may occur even if unions are weak and migration high. An advantage of QCA is that it reveals assumptions not readily apparent in standard Millian comparisons of similarities and differences (Boswell and Brown 1994).

We can estimate a “best model” for the hypothetical case by comparing equations with and without its presence. If we were to assume that the outcome for combination  $MUr$  is  $A$ , then the equation for  $A$  would change from  $A = MR$  to  $A = MR (+ MU)$ . The theory would clearly predict black strikebreaking when high migration is combined with a weak union (irrespective of the state’s attitude). No obvious criteria allow us to rule out this possibility, and the association of outcome  $A$  with  $M$  makes it likely. We enclose the  $MU$  term in parentheses to signify that its inclusion is based on a hypothetical case.<sup>31</sup> Next, assuming that  $MUr = B$  can be reasonably rejected, given that no combination produced outcome  $B$  over the length of the strike. If  $MUr = C$ , then the equation becomes  $C = mU + mR (+ Ur)$ , a possibility that cannot be entirely dismissed if, for instance, a pro-union government were racist. Finally,  $MUr = D$  can be considered logically unlikely. A biracial coalition failed to develop in Milwaukee ( $mUr$ ) despite more favorable circumstances, leaving  $A$  or  $C$  as logical outcomes. Although neither can be ruled out, logical relationships from the historical evidence imply outcome  $A$  as most likely for a comparable northern city. The steel communities were able to use migrant strikebreakers in other cases except Cleveland and Wheeling, where a strong union pursued a biracial strategy. Deriving such implications is important not just for generalizing to other research but also for demonstrating that our empirical findings are not reliant on illogical or unlikely assumptions about hypothetical cases.<sup>32</sup>

<sup>31</sup> Ragin (1987) provides no separate Boolean notation for equation terms based on hypothetical cases. We think using parentheses is consistent with his approach and a useful way of distinguishing hypothetical arguments where additional empirical investigation is needed.

<sup>32</sup> We also considered a five-variable model (recent black migrants, union strength, local government repression, city size, and industry ownership). The 32 possible combinations produced numerous contradictions and hypothetical cases that precluded

## DISCUSSION

The models derived from table 5 support the expected positive relationship between recent migration and black strikebreaking on the one hand but call into question the expected negative relationship between recent migration and the development of biracial labor solidarity in the context of a pro-union local government. The implicit assumption in split labor market theory that solidarity obtains in the absence of recent minority migration should be amended. Further, the equations suggest that the formation of biracial labor coalitions is contingent upon the *sequence* of longer-term factors. Specifically, interracial solidarity may develop only in locations where there is a historically strong union organization. Black strikebreaking develops in response to the more immediate factors of the repressive nature of the local government as well as the presence of recent black migrants. This result is consistent with the fact that instances of interracial solidarity were relatively infrequent during the 1919 strike, suggesting that workers coalesce around class (as opposed to race) interests only when the proper preconditions have been met.

Having tested our results against two alternative specifications, it is useful to return to the original formal model of collective action. In particular, the effects of the local government can be specified and incorporated into an expanded theory of split labor markets. These effects can be modeled indirectly by altering the parameter values: increases of 5% for collective sanctions, individual sanctions, and union monitoring capacity and a 5% decrease in the cost of first-level cooperation. Conversely, we model repressive governments by decreasing collective sanctions, individual sanctions, and monitoring capacity by 5%, while raising the cost of first-level compliance by 5%.

Figures 2 and 3 reflect the revised compliance levels associated with pro-union and anti-union governments, respectively. Here, proportional parameter differences between subgroups are the same as in figure 1, but base levels of  $S_o$ ,  $S_p$ ,  $M$ , and  $K_1$  are altered as described above to reflect either the increased resources of the union and the decreased costs to striking workers where the government is supportive or the decreased resources of the union and the increased costs to workers where the government is repressive. Note that high levels of compliance are likely to emerge only where the state is relatively pro-labor and the union is relatively strong. All eight combinations found in table 5 are now represented by the models plotted in figures 2 and 3.

With the effects of a pro-union government added to the model, both

---

the derivation of equations. In short, the five-variable model added nothing to the more parsimonious version presented in table 5.

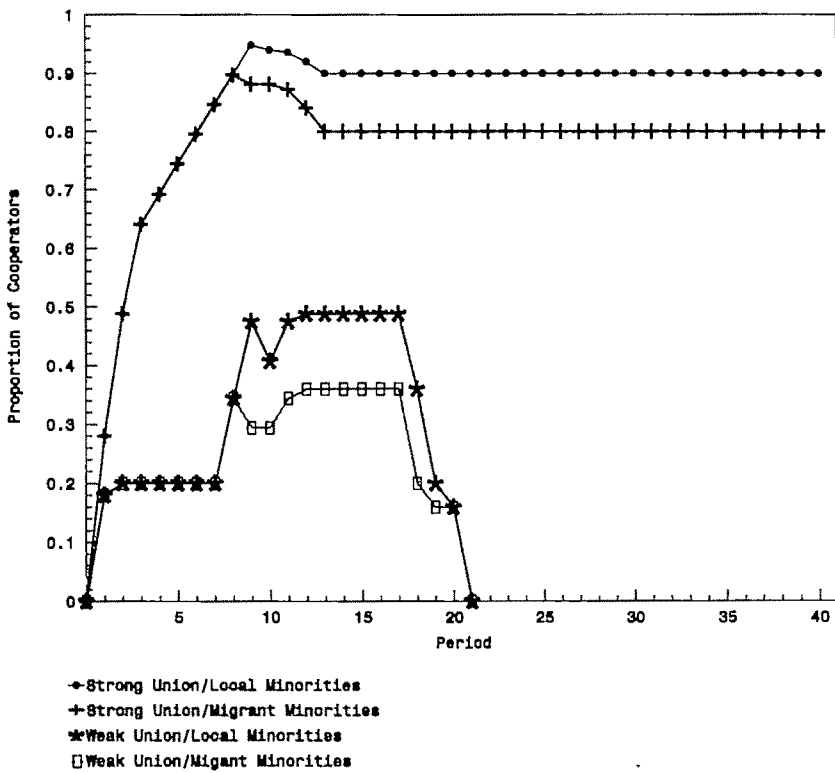


FIG. 2.—Revised model of group compliance levels, pro-union government

strong unions with local minorities and strong unions with migrant minorities achieve high levels of compliance that reach .9 and .8, respectively, once they attain equilibrium. The effects of a pro-union local government have much less dramatic effects where the union is weak. Compared to figure 3, weak unions in figure 2 require several more iterations before the compliance levels fall to zero. This suggests that pro-union governments may increase the viability of collective action where unions are weak, but may not be sufficient to sustain it over the long term, independent of other factors.<sup>33</sup> Additionally, oppositional control does not develop in the pro-union model (full defection becomes

<sup>33</sup> While our model treats the state political context as exogenous, a more refined model might include the state as a corporate actor that changes the structure of the situation (hence the costs, risks, and payoffs) through its decisions. This issue should be considered more fully in future research.

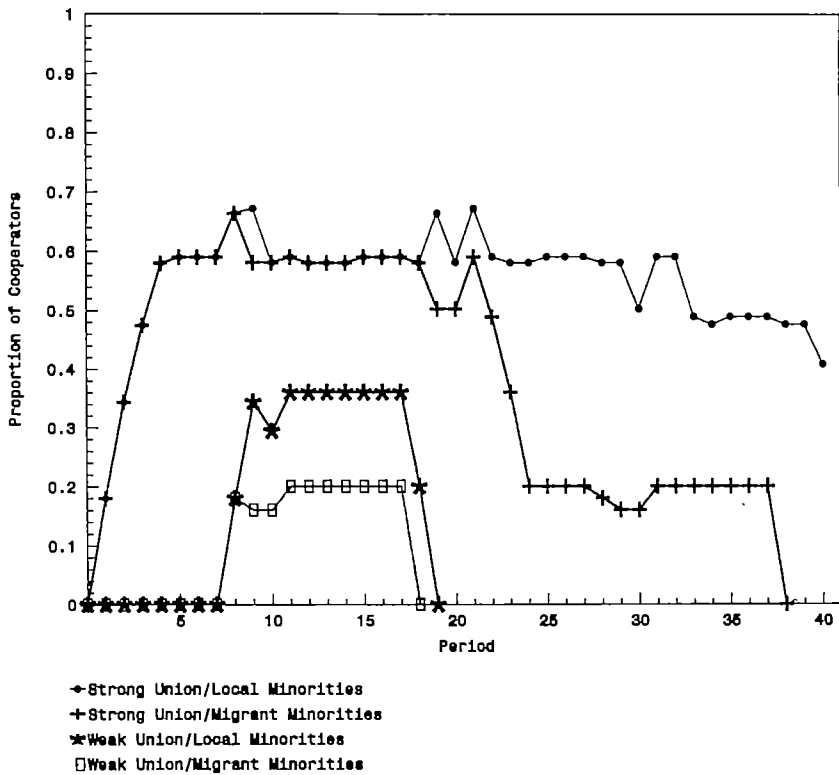


FIG. 3.—Revised model of group compliance levels, anti-union government

the universal strategy in all four cases), whereas oppositional control does develop in three of four anti-union cases.

Our models show that what is important in determining the presence or absence of minority strikebreaking is not necessarily the same as what determines the viability of interracial solidarity. An intuitive argument is that these two conditions are opposites, that the absence of the factors that cause one would result in the other. What our analysis has shown, however, is that the development of biracial labor solidarity requires a favorable political context. This result is congruent with the history of the 1919 strike. Instances of widespread black participation in the strike are rare, suggesting that working-class cohesion is a relatively fragile state requiring a precise combination of class and political factors.

The duration of the strike effort in both Cleveland and Wheeling is indicative of the benefits that accrue to members of interracial labor unions. Although the failure of the strike in other areas meant the even-

tual defeat of all workers, the absence of a racially divided working class contributed to the relative success of the steel workers in these two cities. The presence of a pro-union local government appears to have created conditions favorable to the development of interracial labor organization, ameliorating the negative effects of the split labor market and incoming minority migrants.

## CONCLUSIONS

We have divided our conclusions into their historical, theoretical, and methodological relevance. For understanding the history of the racial conflict during the Great Steel Strike of 1919, we offer a comparative analysis of community differences in the labor markets, union history, and political context. Our measure of recent black migrants with sojourning characteristics documents the extent and variation in the split labor market, which previous research had only assumed. Results from the analysis show that a split labor market was a necessary, although not sufficient, condition for black strikebreaking. Also important to consider is the political context of the strikes. The effect of an influx of minority migrants depended in part on the local government. In the rare cases of biracial labor solidarity, a pro-union government and a strong union were necessary and even sufficient in the face of an influx of black migrants. While we rule out the relevance of city size and differences between the two major steel companies involved in the strike, future historical research might consider other factors such as the ethnic composition of the dominant workers, the degree of racism preceding the migrant influx, or the particular organization of the union local. Expanding the research to other communities or other strikes should also be fruitful.

In terms of split labor market theory, the 1919 strike had been a "crucial test" of the theory, which we subjected to a comparative analysis at the community level. Most important, we expanded the theory to apply to solidarity as well as to strikebreaking by incorporating a formal theory of collective action. Rather than assume strikebreaking as an aberrant and solidarity a passive condition, we postulated that both result from how market and organizational attributes structure the options and interests of workers and unions. We also offer a procedure for measuring aspects of the split that may be useful in related research. The results provide support for our expanded version of the theory.

This support for an expanded split labor market theory is qualified by the salience of variables not directly tied to options and interests in the labor market. In our case, the QCA results were indeterminate until state repression was included. That the pattern of race relations depended heavily on the political context suggests that the articulation of the theory

has been overly economic. While this point has been made before (Burawoy 1981; Boswell 1986), we found that considering market interactions in different contexts incorporates politics into the theory without violating its basic premises. It also encourages the consideration of other contexts. Racist ideologies that skew and bias information, although difficult to measure, are obviously important.

Finally, for comparative historical research in general, the combination of game theory with QCA provides a useful weave between deductive and inductive research. These share the same formal logic, such that the deductive expectations of a formal model can be converted into a truth table for empirical analysis. The QCA approach allows one to add social structure, context, and causal heterogeneity provided by the interpretive case studies to the abstract models, and it exposes contradictions where the theory is inadequate. Historical outcomes and theoretical revisions can then be related back to the formal model for building general theory.

As with any method, ours suffers limitations.<sup>34</sup> A severe problem with QCA is the proliferation of hypothetical cases as variables are added (Griffin et al. 1992, p. 130). Our formal model of collective action truncates this problem by creating a tiered relationship between the interaction of actors with causal attributes and the context that mediates the interaction. To be sure, this does not resolve the problem as multiple contexts or multiple attributes may apply, but it does provide a rationale for ordering their inclusion. Other problems remain, such as the reduction of continuous variation to binary (0/1) categories (Griffin et al. 1992, p. 130). While these limitations require careful consideration, Ragin (1987) points out that these are a modest price for the benefits to comparative historical research. The most important is a formal causal logic that requires consideration of all alternative outcomes while retaining the causal complexity and contextual nuance that is the hallmark of comparative history.

## APPENDIX A

### System Parameters, Variables, and Equations

A complete description of the terms and the derivation of the equations that make up Heckathorn's formal theory of group mediated social control has been provided elsewhere (Heckathorn 1990a, 1990b, 1991). Here, we review the terms and equations to clarify computational procedures and our own modifications.

<sup>34</sup> See *Rationality and Society* (1992) and Ragin (1991) for discussions of issues in game theory and QCA, respectively.

### Parameters

Heckathorn's model of a mixed sanction system is defined by the following fixed parameters:

- $S_c$  Strength of the collective sanction (applied to the actor's group by the agent);  $S_c \geq 0$ .
- $S_i$  Strength of the individual sanction (applied to the actor by the agent);  $S_i \geq 0$ .
- $M$  Agent's monitoring capacity (probability of detecting each first-level defection);  $1 \geq M \geq 0$ .
- $E_2$  Efficacy of actor's level 2 (compliant) control;  $1 \geq E_2 \geq 0$ .
- $E_3$  Efficacy of actor's level 3 (oppositional) control;  $1 \geq E_3 \geq 0$ .
- $K_1$  Cost of level 1 cooperation (costs of complying with the agent).
- $K_2$  Cost of level 2 cooperation (cost of exercising compliant control).
- $K_3$  Cost of level 3 cooperation (cost of exercising oppositional control).
- $N$  Total number of actors in the group.

### Variables

The values of the following variables depend on the prior strategy selections of the actors in the system:

- $N_{c1}$  Number of actors cooperating at level 1 (the number voluntarily complying with the agent);  $0 \leq N_{c1} \leq N - 1$ .
- $N_{c2}$  Number of actors cooperating at level 2 (the number exercising compliant control);  $0 \leq N_{c2} \leq N - 1$ .
- $N_{c3}$  Number of actors cooperating at level 3 (the number exercising oppositional control);  $0 \leq N_{c3} \leq N - 1$ .
- $N_{d1}$  Number of actors choosing to defect at level 1 (the number who choose not to comply with the agent);  $0 \leq N_{d1} \leq N - 1$ .
- $O_{d1}$  Opportunity for group members to defect at level 1 if the actor exercises neither compliant nor oppositional control;  $0 \leq O_{d1} \leq 1$ .
- $O'_{d1}$  Opportunity for group members other than the actor to defect at level 1 reduced by the actor's decision to exercise compliant control;  $0 \leq O'_{d1} \leq 1$ .
- $O''_{d1}$  Opportunity for group members, including the actor, to defect at level 1 increased by



- the actor's exercise of oppositional control;  
 $0 \leq O''_{d1} \leq 1$ .
- $O_{c2}$  Opportunity for group members to cooperate at level 2 by exercising compliant control if the actor does not exercise oppositional control;  $0 \leq O_{c2} \leq 1$ .
- $O''_{c2}$  Opportunity for group members to cooperate at level 2 reduced by the actor's exercise of oppositional control;  $0 \leq O''_{c2} \leq 1$ .
- $R_{cc}, R_{dc}, R_{cd}, R_{co}, R_{do}, R_{dd}$  Risk of sanctions if the actor chooses each of the strategies of full cooperation, hypocritical cooperation, private cooperation, hypocritical opposition, full opposition, and full defection, respectively, where the first subscript refers to the actor's level 1 choice (cooperate/defect) and the second subscript refers to the higher level choice among compliant control (c), oppositional control (o), and higher level defection (d);  $0 \leq R \leq 1$ .
- $P_{cc}, P_{dc}, P_{cd}, P_{co}, P_{do}, P_{dd}$  Payoffs if the actor chooses the strategies full cooperation, hypocritical cooperation, private cooperation, hypocritical opposition, full opposition, and full defection, respectively.

### Equations

We modify Heckathorn's (1990b, p. 5) equations to reflect heterogeneity across subgroups in the following manner:

*Equation set X.*—If actor is a member of subgroup  $N_x$ :

$$\begin{aligned}
 O_{c2} &= [(1 - E_{3x})^{Nc3x}][(1 - E_{3y})^{Nc3y}][(1 - E_{3z})^{Nc3z}] \\
 O''_{c2} &= [(1 - E_{3x})^{(Nc3x+1)}][(1 - E_{3y})^{Nc3y}][(1 - E_{3z})^{Nc3z}] \\
 O_{d1} &= [(1 - E_{2x}O_{c2})^{Nc2x}][(1 - E_{2y}O_{c2})^{Nc2y}][(1 - E_{2z}O_{c2})^{Nc2z}] \\
 O'_{d1} &= [(1 - E_{2x}O_{c2})^{(Nc2x+1)}][(1 - E_{2y}O_{c2})^{Nc2y}][(1 - E_{2z}O_{c2})^{Nc2z}] \\
 O''_{d1} &= [(1 - E_{2x}O''_{c2})^{Nc2x}][(1 - E_{2y}O''_{c2})^{Nc2y}][(1 - E_{2z}O''_{c2})^{Nc2z}] \\
 R_{cc} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O'_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{dc} &= 1 - [(1 - O'_{d1}M_x)^{(Nd1x+1)}][(1 - O'_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{cd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x}][(1 - O_{d1}M_y)^{Nd1y}][(1 - O_{d1}M_z)^{Nd1z}] \\
 R_{co} &= 1 - [(1 - O''_{d1}M_x)^{Nd1x}][(1 - O''_{d1}M_y)^{Nd1y}][(1 - O''_{d1}M_z)^{Nd1z}] \\
 R_{do} &= 1 - [(1 - O''_{d1}M_x)^{(Nd1x+1)}][(1 - O''_{d1}M_y)^{Nd1y}][(1 - O''_{d1}M_z)^{Nd1z}]
 \end{aligned}$$

$$\begin{aligned}
 R_{dd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x+1}][(1 - O_{d1}M_y)^{Nd1y}][(1 - O_{d1}M_z)^{Nd1z}] \\
 P_{cc} &= -S_{cx}R_{cc} - K_{1x} - O_{c2}K_{2x} \\
 P_{dc} &= -S_{cx}R_{dc} - S_{1x}M_xO_{d1} - K_{1x}(1 - O_{d1}) - O_{c2}K_{2x} \\
 P_{cd} &= -S_{cx}R_{cd} - K_{1x} \\
 P_{co} &= -S_{cx}R_{co} - K_{1x} - K_{3x} \\
 P_{do} &= -S_{cx}R_{do} - S_{1x}M_xO'_{d1} - K_{1x}(1 - O'_{d1}) - K_{3x} \\
 P_{dd} &= -S_{cx}R_{dd} - S_{1x}M_xO_{d1} - K_{1x}(1 - O_{d1})
 \end{aligned}$$

*Equation set Y.*—If the actor is a member of subgroup  $N_y$ :

$$\begin{aligned}
 O_{c2} &= [(1 - E_{3x})^{Nc3x}][(1 - E_{3y})^{Nc3y}][(1 - E_{3z})^{Nc3z}] \\
 O'_{c2} &= [(1 - E_{3x})^{Nc3x}][(1 - E_{3y})^{(Nc3y+1)}][(1 - E_{3z})^{Nc3z}] \\
 O_{d1} &= [(1 - E_{2x}O_{c2})^{Nc2x}][(1 - E_{2y}O_{c2})^{Nc2y}][(1 - E_{2z}O_{c2})^{Nc2z}] \\
 O'_{d1} &= [(1 - E_{2x}O_{c2})^{Nc2x}][(1 - E_{2y}O_{c2})^{(Nc2y+1)}][(1 - E_{2z}O_{c2})^{Nc2z}] \\
 O''_{d1} &= [(1 - E_{2x}O''_{c2})^{Nc2x}][(1 - E_{2y}O''_{c2})^{Nc2y}][(1 - E_{2z}O''_{c2})^{Nc2z}] \\
 R_{cc} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O'_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{dc} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O'_{d1}M_y)^{(Nd1y+1)}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{cd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x}][(1 - O_{d1}M_y)^{Nd1y}][(1 - O_{d1}M_z)^{Nd1z}] \\
 R_{co} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O''_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{do} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O''_{d1}M_y)^{(Nd1y+1)}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{dd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x}][(1 - O_{d1}M_y)^{(Nd1y+1)}][(1 - O_{d1}M_z)^{Nd1z}] \\
 P_{cc} &= -S_{cy}R_{cc} - K_{1y} - O_{c2}K_{2y} \\
 P_{dc} &= -S_{cy}R_{dc} - S_{1y}M_yO_{d1} - K_{1y}(1 - O_{d1}) - O_{c2}K_{2y} \\
 P_{cd} &= -S_{cy}R_{cd} - K_{1y} \\
 P_{co} &= -S_{cy}R_{co} - K_{1y} - K_{3y} \\
 P_{do} &= -S_{cy}R_{do} - S_{1y}M_yO'_{d1} - K_{1y}(1 - O'_{d1}) - K_{3y} \\
 P_{dd} &= -S_{cy}R_{dd} - S_{1y}M_yO_{d1} - K_{1y}(1 - O_{d1})
 \end{aligned}$$

*Equation set Z.*—If the actor is a member of subgroup  $N_z$ :

$$\begin{aligned}
 O_{c2} &= [(1 - E_{3x})^{Nc3x}][(1 - E_{3y})^{Nc3y}][(1 - E_{3z})^{Nc3z}] \\
 O'_{c2} &= [(1 - E_{3x})^{Nc3x}][(1 - E_{3y})^{Nc3y}][(1 - E_{3z})^{(Nc3z+1)}] \\
 O_{d1} &= [(1 - E_{2x}O_{c2})^{Nc2x}][(1 - E_{2y}O_{c2})^{Nc2y}][(1 - E_{2z}O_{c2})^{Nc2z}] \\
 O'_{d1} &= [(1 - E_{2x}O_{c2})^{Nc2x}][(1 - E_{2y}O_{c2})^{Nc2y}][(1 - E_{2z}O_{c2})^{(Nc2z+1)}] \\
 O''_{d1} &= [(1 - E_{2x}O''_{c2})^{Nc2x}][(1 - E_{2y}O''_{c2})^{Nc2y}][(1 - E_{2z}O''_{c2})^{Nc2z}] \\
 R_{cc} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O'_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{dc} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O'_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{(Nd1z+1)}] \\
 R_{cd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x}][(1 - O_{d1}M_y)^{Nd1y}][(1 - O_{d1}M_z)^{Nd1z}] \\
 R_{co} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O''_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{Nd1z}] \\
 R_{do} &= 1 - [(1 - O'_{d1}M_x)^{Nd1x}][(1 - O''_{d1}M_y)^{Nd1y}][(1 - O'_{d1}M_z)^{(Nd1z+1)}] \\
 R_{dd} &= 1 - [(1 - O_{d1}M_x)^{Nd1x}][(1 - O_{d1}M_y)^{Nd1y}][(1 - O_{d1}M_z)^{(Nd1z+1)}] \\
 P_{cc} &= -S_{cz}R_{cc} - K_{1z} - O_{c2}K_{2z} \\
 P_{dc} &= -S_{cz}R_{dc} - S_{1z}M_zO_{d1} - K_{1z}(1 - O_{d1}) - O_{c2}K_{2z}
 \end{aligned}$$

$$\begin{aligned}
P_{cd} &= -S_{\alpha}R_{cd} - K_{1s} \\
P_{co} &= -S_{\alpha}R_{co} - K_{1s} - K_{3s} \\
P_{do} &= -S_{\alpha}R_{do} - S_{\alpha}M_{\alpha}O'_{d1} - K_{1s}(1 - O'_{d1}) - K_{3s} \\
P_{dd} &= -S_{\alpha}R_{dd} - S_{\alpha}M_{\alpha}O_{d1} - K_{1s}(1 - O_{d1})
\end{aligned}$$

Calculation of all the opportunities, risks, and payoffs for the appropriate set of equations ( $X$ ,  $Y$ , or  $Z$ ) constitutes one iteration (period) of the model. We then update  $N_{c1}$ ,  $N_{c2}$ ,  $N_{c3}$ , and  $N_{d1}$  based upon the actor's strategy selection and recompute system variables for the next actor's move at the following iteration.

## APPENDIX B

### Migration Data

TABLE B1  
INDICATORS OF RECENT MIGRATION AMONG BLACK MALES 20-44 YEARS OLD,  
1910-20

City	%, 1920	Relative Rate of Increase, 1910-20	% Increase, 1910-20	Migrant†
Bethlehem, Penn. ....	.80	- 18	200	0
Buffalo, N.Y. . ....	1 47	.07	188	1
Chicago .....	5 43	-.01	137	1
Cleveland . . ....	5 91	.01	337	1
Decatur, Ill. . . ....	1.59	- 47	60	0
East Chicago, Ind. ...	5 62	.06	4,308	1
Gary, Ind. . . . .	10 58	10	968	1
Johnstown, Penn. ....	5.64	41	494	1
Joilet, Ill. ....	2 40	.01	32	0
McKeesport, Penn.	2.36	.04	18	0
Milwaukee .....	.73	- 07	143	0
New Castle, Penn. ....	2 65	.13	71	0
Pittsburgh .... .	8.38	.03	43	1
Reading, Penn. ....	1 17	.05	18	0
Wheeling, W.V. ...	3.99	.06	36	0
Youngstown, Ohio ...	7 28	.18	290	1

SOURCES — U S. Bureau of the Census (1913, 1918, 1922)

\* Black males 20-44 years old as a percentage of all males 20-44 years old

† Coded "1" if at least two of the following three criteria are met col. 1  $\geq 5.00$ , col. 2  $\geq 0.00$ ; col. 3  $\geq 100$

TABLE B2

INDICATORS OF RECENT MIGRATION AMONG BLACK SINGLE MALES, 1910-20

City	%*, 1920	Relative Rate of Increase, 1910-20	% Increase, 1910-20	Migrant†
Bethlehem, Penn. ....	1.21	-.34‡	196‡	0
Buffalo, N Y .....	1.33	.07	143	1
Chicago . . . . .	4.46	.00	112	1
Cleveland .....	5.33	.04	309	1
Decatur, Ill. ....	2.74	-.47	15	0
East Chicago, Ind. ....	5.95	.09‡	4,428	1
Gary, Ind. ...	9.71	-.03‡	967	1
Johnstown, Penn. ....	7.96	.18	682	1
Joliet, Ill. ....	2.18	-.02	28	0
McKeesport, Penn. ....	2.20	.04	44	0
Milwaukee .....	.64	-.01	118	0
New Castle, Penn. ....	2.84	.02	56	0
Pittsburgh ..	7.07	.02	37	1
Reading, Penn. ....	1.27	.47	48	0
Wheeling, W V. ....	3.63	.11	36	0
Youngstown, Ohio . ...	7.50	.12	269	1

SOURCES.—U.S. Bureau of the Census (1913, 1918, 1922).

\* Single black males as a percentage of all single males.

† Coded "1" if at least two of the following three criteria are met: col. 1  $\geq 5.00$ , col. 2  $\geq 0.00$ , col. 3  $\geq 100$ .

‡ Estimate derived using the ratio of single black males to total black males for 1920, exact figures for 1910 not available.

## REFERENCES

- Asher, Robert. 1978. "Painful Memories: The Historical Consciousness of Steelworkers and the Steel Strike of 1919." *Pennsylvania History* 45:61-86.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic.
- . 1986. "An Evolutionary Approach to Norms." *American Political Science Review* 80:1097-1111.
- Berg, Elliot J. 1961. "Backward Sloping Supply Functions in Dual Economies: The African Case." *Quarterly Journal of Economics* 75:468-92.
- Bianco, William T., Peter C. Ordeshook, and George Tsebelis. 1990. "Crime and Punishment: Are One-Shot, Two-Person Games Enough?" *American Political Science Review* 84:569-86.
- Binmore, Ken, and Partha Dasgupta. 1986. "Game Theory: A Survey." Pp. 1-45 in *Economic Organizations as Games*, edited by K. Binmore and P. Dasgupta. New York: Basil Blackwell.
- "The Black Man and the Unions." 1918. *Crisis* 16:216-17.
- Bonacich, Edna. 1972. "A Theory of Ethnic Antagonism: The Split Labor Market." *American Sociological Review* 37:547-59.
- . 1975. "Abolition, the Extension of Slavery, and the Position of Free Blacks: A Study of Split Labor Markets in the United States, 1830-1863." *American Journal of Sociology* 81:601-28.
- . 1976. "Advanced Capitalism and Black/White Race Relations in the United States: A Split Labor Market Interpretation." *American Sociological Review* 41:34-51.

- . 1979. "The Past, Present, and Future of Split Labor Market Theory." *Research in Race and Ethnic Relations* 1:17–64.
- Bonacich, Edna, and Lucie Cheng. 1984. "A Theoretical Orientation to International Labor Migration." Pp. 1–56 in *Labor Migration under Capitalism*, edited by L. Cheng and E. Bonacich. Berkeley and Los Angeles: University of California Press.
- Boryczka, Raymond, and Lorin Lee Cary. 1982. *No Strength without Union: An Illustrated History of Ohio Workers, 1803–1980*. Columbus: Ohio Historical Society.
- Boswell, Terry E. 1986. "A Split Labor Market Interpretation of Discrimination against Chinese Immigrants, 1850–1882." *American Sociological Review* 51:352–71.
- Boswell, Terry, and Cliff Brown. 1994. "The Scope of General Theory: Methods for Linking Deductive and Inductive Comparative History." Manuscript. Emory University, Department of Sociology.
- Boswell, Terry, and David Jorjani. 1988. "Uneven Development and the Origins of Split Labor Market Discrimination: A Comparison of Black, Chinese, and Mexican Immigrant Minorities in the United States." Pp. 169–85 in *Racism, Sexism, and the World System*, edited by J. Smith, J. Collins, T. K. Hopkins, and A. Muhammad. New York: Greenwood.
- Brody, David. 1965. *Labor in Crisis: The Steel Strike of 1919*. Philadelphia: Lippincott.
- Brown, Cliff, and Terry Boswell. 1994. "Modeling Racially Split Labor Markets: Theoretical Adaptations and Technical Modifications for Heckathorn's Sequential Decision Model." Manuscript. Emory University, Department of Sociology.
- Burawoy, Michael. 1981. "The Capitalist State in South Africa: Marxist and Sociological Perspectives on Race and Class." Pp. 279–335 in *Political Power and Social Theory*, edited by Maurice Zeitlin. Greenwich, Conn.: JAI Press.
- Dickerson, Dennis D. 1986. *Out of the Crucible: Black Steelworkers in Western Pennsylvania*. Albany: State University of New York Press.
- Donald, Henderson H. 1921. "The Negro Migration of 1916–1918." *Journal of Negro History* 6:383–499.
- Drass, Kriss, and Charles Ragin. 1986. *QCA: A Microcomputer Package for Qualitative Comparative Analysis of Social Data*. Evanston, Ill.: Center for Urban Affairs and Policy Research, Northwestern University.
- Eckstein, Harry. 1975. "Case Study and Theory in Political Science." Pp. 79–137 in *Handbook of Political Science*. Vol. 7, *Strategies of Inquiry*, edited by Fred I. Greenstein and Nelson Polsby. Reading, Mass.: Addison-Wesley.
- Edwards, Richard. 1979. *Contested Terrain*. New York: Basic.
- Epstein, Abraham. 1918. *The Negro Migrant in Pittsburgh*. Pittsburgh: University of Pittsburgh Press.
- Fligstein, Neil. 1981. *Going North: Migration of Blacks and Whites from the South, 1900–1950*. New York: Academic Press.
- Florant, Lyonel C. 1942. "Negro Internal Migration." *American Sociological Review* 7:782–91.
- Foner, Philip S. 1973. *History of the Labor Movement in the United States*, vol. 3. New York: International Publishers.
- . 1974. *Organized Labor and the Black Worker*. New York: Praeger.
- . 1988. *History of the Labor Movement in the United States*, vol. 8. New York: International Publishers.
- Foster, William Z. 1920. *The Great Steel Strike and Its Lessons*. New York: Huebsch.
- Gadsby, M. A. 1919. "The Steel Strike." *Monthly Labor Review* 9:79–94.
- Garraty, John A. 1960. "The United States Steel Corporation versus Labor: The Early Years." *Labor History* 1:3–38.
- Gottlieb, Peter. 1978. "Migration and Jobs: The New Black Workers in Pittsburgh, 1916–30." *Western Pennsylvania Historical Magazine* 61:1–15.

- . 1987. *Making Their Own Way: Southern Blacks' Migration to Pittsburgh, 1916–30*. Urbana: University of Illinois Press.
- Griffin, Larry J., Christopher Botsko, Ana-Maria Wahl, and Larry W. Isaac. 1992. "Theoretical Generality, Case Particularity: Qualitative Comparative Analysis of Trade Union Growth and Decline." *International Journal of Comparative Sociology* 32:110–36.
- Grossman, James R. 1989. *Land of Hope: Chicago, Black Southerners, and the Great Migration*. Chicago: University of Chicago Press.
- Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley and Los Angeles: University of California Press.
- Heckathorn, Douglas D. 1988. "Collective Sanctions and the Creation of Prisoner's Dilemma Norms." *American Journal of Sociology* 94:535–62.
- . 1989. "Collective Action and the Second-Order Free-Rider Problem." *Rationality and Society* 1:78–100.
- . 1990a. "Collective Sanctions and Compliance Norms: A Formal Theory of Group-Mediated Social Control." *American Sociological Review* 55:366–84.
- . 1990b. "A Sequential Decision Model for Mixed Sanction Systems." *Policy Studies Research Group*. Working Paper no. 125. University of Missouri—Kansas City, Department of Sociology.
- . 1991. "Collective Sanctions and Group Heterogeneity: Cohesion and Polarization in Normative Systems." *Advances in Group Processes* 9:41–63.
- . 1993. "Collective Action and Group Heterogeneity: Voluntary Provision versus Selective Incentives." *American Sociological Review* 58:329–50.
- Henri, Florette. 1975. *Black Migration: Movement North, 1900–1920*. Garden City, N.Y.: Anchor/Doubleday.
- Interchurch World Movement. 1920. *Report on the Steel Strike of 1919*. New York: Harcourt, Brace, & Howe.
- Jones, Eugene Kinckle. 1919. "The Negro in Industry." *Proceeding of the National Conference of Social Work*. Chicago: University of Chicago Press.
- Kennedy, Louise V. 1930. *The Negro Peasant Turns Cityward: Effects of Recent Migrations to Northern Centers*. New York: Columbia University Press.
- Korpi, Walter, and Michael Shalev. 1979. "Strikes, Industrial Relations, and Class Conflict in Capitalist Societies." *British Journal of Sociology* 30:164–67.
- Kreps, David M. 1990. *Game Theory and Economic Modelling*. Oxford: Clarendon.
- Lemann, Nicholas. 1991. *The Promised Land: The Great Black Migration and How It Changed America*. New York: Knopf.
- Lewis, Edward E. 1931. *The Mobility of the Negro: A Study in the American Labor Supply*. New York: Columbia University Press.
- Lieberson, Stanley. 1991. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70:307–320.
- Lindbeck, Assar. 1993. *Unemployment and Macro-Economics*. London: MIT Press.
- Marks, Carole. 1981. "Split Labor Markets and Black-White Relations, 1865–1920." *Phylon* 43:293–308.
- . 1983. "Lines of Communication, Recruitment Mechanisms, and the Great Migration of 1916–1918." *Social Problems* 31:73–83.
- McAdam, Doug. 1986. "Recruitment of High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92:64–90.
- Murray, Robert K. 1951. "Communism and the Great Steel Strike of 1919." *Mississippi Valley Historical Review* 38 445–66.
- . 1955. *Red Scare: A Study in National Hysteria, 1919–1920*. Minneapolis: University of Minnesota Press.
- Myrdal, Gunnar. 1944. *An American Dilemma: The Negro Problem and Modern Democracy*. New York: Harper.

- Nash, John. 1951. "Non-Cooperative Games." *Annals of Mathematics* 54:286-95.
- National Urban League. 1930. *Negro Membership in American Labor Unions*. New York: Alexander Press.
- Newell, Barbara Warne. 1961. *Chicago and the Labor Movement: Metropolitan Unionism in the 1930's*. Urbana: University of Illinois Press.
- Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 85:1356-75.
- Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, Mass.: Harvard University Press.
- Peters, Ralph. 1994. "Social Perpetuity over Social Transformation: The Failure of Bi-racial Textile Unionization in the Postwar South." Dissertation. Emory University, Department of Sociology.
- Peterson, Florence. 1937. *Strikes in the United States: 1880-1936*. Washington: U.S. Government Printing Office.
- Piore, Michael J. 1979. *Birds of Passage: Migrant Labor and Industrial Societies*. New York: Cambridge University Press.
- Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley-Interscience.
- Ragin, Charles C. 1987. *The Comparative Method*. Berkeley and Los Angeles: University of California Press.
- , ed. 1991. *Issues and Alternatives in Comparative Social Research*. New York: Brill.
- Ragin, Charles C., Susan E. Mayer, and Kriss A. Drass. 1984. "Assessing Discrimination: A Boolean Approach." *American Sociological Review* 49:221-34.
- Rationality and Society*. 1992. Vol. 4, no. 1.
- Reich, Michael. 1981. *Racial Inequality: A Political-Economic Analysis*. Princeton, N.J.: Princeton University Press.
- Scott, Emmett J. 1920. *Negro Migration during the War*. New York: Oxford University Press.
- Sewell, William H., Jr. In press. "Three Temporalities. Toward Eventful Sociology." In *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Siu, Paul C. P. 1952. "The Sojourner." *American Journal of Sociology* 58:34-44.
- "Southern Negro in Cleveland Industries." 1924. *Monthly Labor Review* 19:41-44.
- Spero, Sterling D., and Abram L. Harris. [1931] 1969. *The Black Worker: The Negro and the Labor Movement*. New York: Atheneum.
- Stinchcombe, Arthur. 1990. *Information and Organizations*. Berkeley and Los Angeles: University of California Press.
- Taylor, Michael. 1987. *The Possibility of Cooperation*. Cambridge: Cambridge University Press.
- Trotter, Joe William. 1985. *Black Milwaukee: The Making of an Industrial Proletariat, 1916-45*. Urbana: University of Illinois Press.
- U.S. Bureau of the Census. 1913. *Thirteenth Census of the United States*, vols. 2-3. Washington, D.C.: Government Printing Office.
- . 1918. *Negro Population, 1790-1915*. Washington, D.C.: Government Printing Office.
- . 1922. *Fourteenth Census of the United States*, vols. 2-3. Washington, D.C.: Government Printing Office.
- Whitney, Edson L. 1920. "Strikes and Lockouts in 1916, 1917, 1918, and 1919." *Monthly Labor Review* 10:199-218.
- Wilson, William J. 1978. *The Declining Significance of Race*. Chicago: University of Chicago Press.
- Yellen, Samuel. 1974. *American Labor Struggles, 1877-1974*. New York: Monad Press.

# **SYMPOSIUM ON PREDICTION IN THE SOCIAL SCIENCES**

## **Introduction: Reflections on Historical Prophecy in the Social Sciences<sup>1</sup>**

Michael Hechter  
*University of Oxford*

The collapse of communist regimes in Russia and Eastern Europe was greeted by near-universal enthusiasm, save among two quite different groups. The first, not surprisingly, consists of the former officials of these regimes—many of whom feared for their careers, if not for their very lives, in the wake of lustration. Little of their disappointment was aired in public; it would, no doubt, have fallen on deaf ears. Yet since former government officials probably benefited disproportionately from the opportunities afforded by the new order, many of them may not have been so terribly disappointed anyway.

The second group, more surprisingly, consists of social scientists who failed to anticipate the collapse of communism despite claims to the predictive status of their trade. Successful prediction resolves uncertainty about future states of the world. On almost any view of human nature, certainty is preferred to uncertainty (Friedman, Hechter, and Kanazawa 1994). All of us would like to know what our spouse will be like 10 years after marriage, what successes and failures will befall us in one career as against another, and what kind of a house will have the greatest appreciation 25 years hence. Sociology is no less committed to prediction than are its sister disciplines: from its very beginnings, the attainment of predictive knowledge was one of the field's principal rationales (Schuessler 1968).<sup>2</sup> Perhaps this explains why social scientists have engaged in a

<sup>1</sup> Address correspondence to Michael Hechter, Department of Applied Social Studies, University of Oxford, Barnett House, Wellington Square, Oxford OX1 2BN, United Kingdom.

<sup>2</sup> Thus for Collins (in this issue), "The ability of sociology to make valid predictions is a sign of the maturity of the discipline." There are, of course, other traditions in sociology that are skeptical about the field's capacity to generate predictive knowledge.



good deal more public hand-wringing than the deposed former communist officials. Indeed, two successive program committees (1992, 1993) of the American Sociological Association devoted thematic panels to explore why the dramatic events of 1989 came as such a surprise (Lipset and Bence 1994).

The symposium presented here is derived from the most recent of these panels.<sup>3</sup> Its central question is whether events such as those of 1989 should—or could—have been predicted. More specifically, are social scientists failing to live up to their promises, or do they uphold unrealistic expectations by assuming that phenomena such as revolutions and the collapse of empires in fact are predictable?

That some kinds of events can be predicted with precision is undeniable. The tides and celestial appearances, such as Halley's comet and solar eclipses, have been predicted with admirable precision for centuries. Even more impressive, the existence of difficult-to-observe phenomena, such as antimatter and black holes that were never even imagined but were predicted on the basis of theories like general relativity, have been confirmed occasionally in the history of science.

Of course, many other physical phenomena cannot be predicted nearly as precisely nor as far in advance. Earthquakes provide an instructive example. Since they are one of the most serious types of natural disasters, accurate earthquake prediction could save many lives. The accumulation of stress and the weakening of rock that precede an earthquake have measurable consequences. The sudden lowering of groundwater levels, the existence of tilts and bulges in the earth's surface, changes in the velocity of propagation of *P* and *S* waves, changes in the earth's magnetic field, and increased concentrations of rare gases in well water have been observed in different combinations prior to some—but not all—earthquakes. These observations can be used to signal the possible appearance of an earthquake, but they do not always foretell one.

Seismologists can predict the locations of earthquakes very precisely; their predictions of the intensity of tremors are less accurate. Despite this, cities at risk of earthquakes can adopt building codes to minimize mortality rates. Yet temporal predictions of earthquakes could save even more lives; if predictions were accurate, then potential victims could be evacuated and would survive. Unfortunately, errors in predicting the timing of earthquakes average 20 years or more. It is likely that earthquake predictions will improve with better data. Seismologists have to interpret tectonic dynamics that have taken place for tens of thousands

<sup>3</sup> The first was published as a special issue of *Theory and Society* (Hechter and Szelenyi 1994).

of years on the basis of little more than a century of observation. Similarly, weather forecasts have improved on the basis of new technologies of measurement and computation.

Since revolutions and civil war also cause mortality, their potential victims have a strong interest in accurate predictions of these events. But are social phenomena predictable? Indeed some of them are. Selling automobile insurance can be profitable because statistical data permit insurers to estimate the risk of loss due to crime or accident for various classes of policyholders (Heimer 1985). Economists can predict the effects of different rules of auction on market efficiency (Smith 1991, pp. 25–29). Sociologists can predict the constancy of inequalities in educational attainment under a wide range of conditions (Raftery and Hout 1993; Shavit and Blossfeld 1993). But many other kinds of social phenomena do not seem very amenable to prediction. Thus people cannot easily predict their own future preferences (Kahneman and Snell 1993). Economists cannot predict fluctuations in the stock market (but see Rappoport and White 1994), and their recent macroeconomic forecasts have inspired little confidence in any quarter.

What accounts for the differential ability to predict future events? The standard answer to this question was given by Karl Popper (1957, 1963), who distinguished between two different types of predictions: scientific predictions and unconditional historical prophecies. Most scientific predictions are conditional: they assert that certain changes in the parameters of a given system will be associated with changes in outcomes. So if the price of a given commodity is doubled we can predict that less of it will be demanded. Unconditional scientific predictions can sometimes be derived from these conditional scientific predictions, together with historical statements that assert that the conditions in question are fulfilled.

In contrast, long-term historical prophecies can be derived from scientific conditional predictions only if they apply to systems that are *well-isolated, stationary, and recurrent*. The reason we can make long-term predictions of solar eclipses is because the solar system is a stationary and repetitive system. It is so because it happens to be isolated from the influence of other mechanical systems by vast amounts of empty space. Therefore it is relatively free of potentially destabilizing interference from outside. Recurrent systems of this kind also exist in biology. Systems of this kind occur rarely, however, and there is no reason to expect that long-term unconditional prophecy can be applied to human history. This is because societies can change—in some cases, rather rapidly—due to the growth of technology and other forms of human knowledge. More generally, the problem is that “we cannot predict, scientifically, results which we shall obtain in the course of our own knowledge” (Popper

1982, p. 62). On this account, social situations different from anything that ever happened before can and often do arise, thereby vitiating any possibility for long-term prophecy.<sup>4</sup>

This verdict need not be regarded as terribly pessimistic, for the utility of the physical sciences may not rest in their ability to accurately forecast events like solar eclipses. Likewise for the social sciences. Instead, the main task of the theoretical social sciences, according to Popper, is to trace the unintended social consequences of intentional human actions. This is not so different from the situation of the natural sciences. Both lead to propositions stating what cannot be done. The second law of thermodynamics informs us that it is impossible to build a machine that is perfectly efficient. Likewise, the law of demand informs us that it is impossible to raise the price of a given commodity and have more people willing to buy it. The social sciences provide us with an idea of what can, and what cannot, be done in the sphere of public policy. The role of science in social life is "the modest one of helping us to understand even the more remote consequences of possible actions, and thus of helping us to choose our actions more wisely" (Popper 1963, p. 343).

The conventional wisdom therefore suggests that social phenomena like revolutions are not predictable with any high degree of precision. If so, it follows that social scientists should not hang their heads in shame for failing to anticipate the collapse of the Soviet Union in the year 1989.

For somewhat different reasons, both Kuran and Tilly endorse this conventional wisdom. Kuran's analysis of the inherent unpredictability of revolution hinges on the disjuncture between public and private preferences. Public preferences are opinions that citizens are willing to reveal to survey interviewers and political authorities, but private preferences are not so readily revealed. Revolutionary situations, for Kuran, occur when existing regimes lose legitimacy and citizens desire to replace their leaders, if not the entire system of governance. Fear of sanctions, however, causes these preferences to remain private, and so they are only imperfectly observable. Hence a society might be on the verge of a revolu-

<sup>4</sup> Popper certainly was not alone in arguing for the indeterminacy of history. Morgenstern (1928) also took a pessimistic view of long-term historical prediction, and for related reasons. Once the prediction of an event was made public, Morgenstern averred, people might alter their behavior intentionally so as to then invalidate the prediction. In another early contribution, Merton (1936) made precisely the converse observation that some public prophecies may come to pass precisely because they are self-fulfilling. By highlighting the self-altering potential of historical predictions, both Morgenstern and Merton underline Popper's contention about the destabilizing role of knowledge in complex systems. The recent popularity of the discourse of path dependence among historians presumably attests to similar concerns. Grunberg and Modigliani (1954) and Simon (1954) claim to have refuted Morgenstern's conclusion, but both papers subsequently have been challenged (Henshel 1994).

tionary upheaval without anyone knowing it. In this situation a small, intrinsically insignificant event might be enough to topple the entire social order. Informational limits and observability loom large in Kuran's analysis. Due to the interdependence of individual preferences, revolution always is the outcome of a nonlinear process. This nonlinearity allows for large effects in the consequences of a given event. Imperfect observability of the interdependencies among public preferences means that potentially destabilizing events can never be accurately forecast—hence, the ease of explaining a revolution post hoc and the difficulty of predicting one *ex ante*.

Yet some revolutions will be more predictable than others. This is because private preferences are more difficult to measure and interpret in nondemocratic than in democratic regimes. Thus, due to public opinion surveys, private preferences are generally more observable in democracies. Even so, responses to questions on certain highly charged issues are unlikely to be accurate. To some degree, systematic response bias can be overcome by the use of new techniques in survey analysis. In contrast, nondemocratic regimes both discourage truthful revelation of public preferences and prohibit the collection and dissemination of accurate public opinion data. Sensitive observers should be able to make use of the distinction between public and private preferences, however, to generate hunches about the legitimacy of any given regime. Kuran concludes by predicting that no theory will be able to predict future revolutions accurately. The strength of this claim rests on the interpretation of accuracy. It is not surprising that Kuran takes a stringent view of the matter: an accurate prediction will specify that a revolution will occur in a specific state within a five-year confidence interval at maximum.

Tilly likewise is skeptical of our ability to make exact predictions about macrosocial events. Yet he seems to regard himself a voice in the wilderness. In his highly informed opinion, most current macrosociologists carry out their research as if it were possible to make accurate predictions. This is due to the (mostly implicit) assumptions they make about invariant units of analysis and social processes. More specifically, macrosociologists tend to assume coherent, durable self-propelling social units that are characterized by some general condition. An invariant model of that condition or process is invoked or invented, which enables them to explain the unit's behavior as consistent with that of the invariant model. Citing chapter and verse from a number of currently influential studies, Tilly argues that the so-called predictions deriving from this method tend merely to be definitional and tautological. His acidic description of the activity he terms "improving the model" should be required reading for far too many students and practitioners of sociological research. In this way Tilly adds his voice to others who recently have surveyed compara-

tive-historical methodology and found much of it wanting (Kiser and Hechter 1991; Lieberman 1991; Goldthorpe 1994).

Far from denying the importance of transhistorical regularities in political life, Tilly insists that they do not operate in the form of recurrent structures and processes on a large scale. Whereas students of revolution have assumed they were studying phenomena like tides or celestial appearances, in reality they are studying phenomena like floods. Although floods are natural phenomena that are consistent with general principles like the physics of incompressible fluids in open channels, they cannot be precisely predicted because every instance is different. Time, place, and sequence strongly influence how the relevant processes unfold. Tilly concludes by sketching an alternative way of thinking about macrosocial events and by describing some recent studies that break with past practice in ways he regards as hopeful. Whether or not Tilly's diagnosis and cure is accepted, it is remarkable that it needs to be repeated in 1995. After all Popper, who hardly was an obscure figure in the social sciences, made substantially the same point at least four decades ago.<sup>5</sup>

Perhaps a clue about why some social scientists have turned their backs on this conventional wisdom may be gleaned from Collins's article. This is because Collins begins by asserting not only that it is in principle possible to predict revolutions, but that he actually did so with respect to the demise of the Soviet Union. Collins documents this claim in his article and describes in some detail how he made this particular prediction. Was Collins's anticipation of the Soviet demise an accurate prediction derived from a set of fundamental theoretical principles, or was it merely a lucky guess?

Collins claims that his prediction was based on a geopolitical theory that links a variety of state resources and interstate competition with shifts in legitimacy. His elaborate discussion of the theory constitutes the substantive heart of the article. In a nutshell, geopolitical theory makes a set of conditional predictions about the power and prestige of states. State power and prestige, in turn, are held to vary directly with state

<sup>5</sup> "History is characterized by its interest in actual, singular, or specific events, rather than in laws or generalizations. This view is perfectly compatible with the analysis of scientific method, and especially of causal explanation. . . . The situation is simply this: while the theoretical sciences are mainly interested in finding and testing universal laws, the historical sciences take all kinds of universal laws for granted and are mainly interested in finding and testing singular statements. . . . All causal explanations of a singular event can be said to be historical insofar as the 'cause' is always described by singular initial conditions. And this agrees entirely with the popular idea that to explain a thing causally is to explain it and why it happened, that is to say, to tell its 'story.' But it is only in history that we are really interested in the causal explanation of a *singular* event. In the theoretical sciences, such causal explanations are mainly means to a different end—the testing of universal laws" (Popper 1957, p. 143).

legitimacy (which, for him, is the proximal cause of revolution). Hence, when the power and prestige of states decrease below some threshold level, state breakdown will occur. Analytically, this sounds very much like Popper's description of conditional scientific prediction. To this point, therefore, nothing in Collins's article violates the conventional wisdom in the philosophy of science.

Despite occasional rhetorical flourishes, Collins actually makes no strong claims for the predictive accuracy of geopolitical theory, at least with respect to temporality. He estimates that its predictions have an error term of from 30 to 50 years. In Kuran's view, this performance is unacceptably imprecise. There are at least two reasons for the relative indeterminacy of geopolitical theory. First, each application of the theory requires a specification of initial conditions, which are historically variable. The second reason for the theory's indeterminacy is captured in Collins's discussion of three kinds of historical time: the *longue durée*, in which geopolitical resources change slowly, wartime, in which large changes can occur within a small number of years, and social movement time, the two or three days during which the legitimacy of the threatened state—and hence its very fate—hangs in the balance. The implication is that events in the *longue durée* are highly determinate, those during wartime are less so, and those during social movement time are essentially indeterminate. Since state breakdown is a function of all three kinds of time, Collins therefore concedes Kuran's point that revolutions cannot be predicted with any precision. And he would appear to agree with Tilly that revolutions are not invariant processes.

Altogether, then, this symposium suggests that our situation with respect to prediction is akin to that in seismology. There is reason to believe that it is possible to predict the location of major social upheavals. Predicting their intensity is more dubious. Predicting their timing is likely to be beyond our grasp both now and in the future. Even so, Collins ends his article on a distinctly upbeat note. Readers must decide for themselves if this difference in tone between Collins, on the one hand, and Kuran and Tilly, on the other, is semantic—that is, whether it amounts to a dispute between those who view the glass as half full or half empty—or if something more fundamental is at stake.

## REFERENCES

- Friedman, Debra, Michael Hechter, and Satoshi Kanazawa. 1994. "A Theory of the Value of Children." *Demography* 31 (3): 375–401.
- Goldthorpe, John H. 1994. "The Uses of History in Sociology—a Reply." *British Journal of Sociology* 45 (1): 55–77.
- Grünberg, E., and F. Modigliani. 1954. "The Predictability of Social Events." *Journal of Political Economy* 62:465–478.

- Hechter, Michael, and Ivan Szelenyi, eds. 1994. "Theoretical Implications of the Demise of State Socialism." Special issue of *Theory and Society* 23 (2).
- Heimer, Carol. 1985. *Reactive Risk and Rational Action*. Berkeley and Los Angeles: University of California Press.
- Henshel, Richard. 1994. "Critique of the Grünberg and Modigliani and Simon Possibility Theorem for Self-Altering Predictions." Paper presented at the Thirteenth World Congress of Sociology, International Sociological Association. Bielefeld, Germany, July.
- Kahneman, Daniel, and Jackie S. Snell. 1993. "Predicting a Changing Taste: Do People Know What They Will Like?" *Journal of Behavioral Decision Making* 5 (3): 187-200.
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97 (1): 1-30.
- Lieberson, Stanley. 1991. "Small *N*'s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70 (2): 307-20.
- Lipset, Seymour M., and Gyorgyi Bence. 1994. "Anticipations of the Failure of Communism." *Theory and Society* 23 (2): 169-210.
- Merton, Robert K. 1936. "The Unanticipated Consequences of Purposive Social Action." *American Sociological Review* 1:894-904.
- Morgenstern, Oskar. 1928. *Wirtschaftsprognose—eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Vienna: Springer.
- Popper, Karl R. 1957. *The Poverty of Historicism*. London: Routledge & Kegan Paul.
- . 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge & Kegan Paul.
- . 1982. *The Open Universe: An Argument for Indeterminism*. London: Routledge.
- Raftery, Adrian E., and Michael Hout. 1993. "Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75." *Sociology of Education* 66 (1): 41-62.
- Rappoport, Peter, and Eugene N. White. 1994. "Was the Crash of 1929 Expected?" *American Economic Review* 84 (1): 271-81.
- Schuessler, Karl F. 1968. "Prediction." Pp. 418-25 in *International Encyclopedia of the Social Sciences*, vol. 11. Edited by David L. Sills. New York: Macmillan.
- Shavit, Yossi, and Hans Peter Blossfeld. 1993. *Persistent Inequality*. Boulder, Colo.: Westview Press.
- Simon, Herbert A. 1954. "Bandwagon and Underdog Effects and the Possibility of Election Predictions." *Public Opinion Quarterly* 18:245-53.
- Smith, Vernon L. 1991. *Papers in Experimental Economics*. Cambridge: Cambridge University Press.

# The Inevitability of Future Revolutionary Surprises<sup>1</sup>

Timur Kuran  
*University of Southern California*

Although social scientists have written extensively on revolutions, none of their theories has shown much predictive success in practice. Recent revolutions surprised social scientists as much as anyone else. This article proposes that revolutionary surprises will occur repeatedly, although it is possible to identify countries relatively likely to experience a sudden explosion. The argument hinges on preference falsification—the act of misrepresenting one's preferences under perceived social pressures. By falsifying their preferences with regard to the incumbent regime, disgruntled citizens distort perceptions of the potential for political change. The article's key proposition may be refuted by building a model that successfully predicts when and where revolutions will occur.

## THE FALL OF COMMUNISM AND OTHER UNANTICIPATED OVERTURNS

Intellectuals disagree about many things, so there is nothing unusual about the numerous controversies that have followed the fall of East European communism. What is remarkable is our nearly unanimous agreement on the fact that this momentous overturn caught the world by surprise. The evidence is overwhelming that virtually no one expected communism to collapse rapidly, with little bloodshed, and throughout Eastern Europe before the end of the 1980s. The stunned may be grouped in four categories.<sup>2</sup>

One category consists of observers from outside the region, including journalists, diplomats, statesmen, futurologists, and scholars. Within the

<sup>1</sup> I presented an earlier version of this paper at a session entitled "Why Didn't We See It Coming? The Case of the Soviet Collapse," held in Miami at the August 1993 convention of the American Sociological Association. I benefited from the comments of the session's discussant, Edgar Kiser, and also from those of two *AJS* reviewers. Correspondence should be directed to Timur Kuran, Department of Economics, University of Southern California, Los Angeles, California 90089-0253.

<sup>2</sup> Further details on the information provided in this section may be found in Kuran (1991). See also Ash (1990) and Tismaneanu (1992).



last group, renowned East European specialists were caught by surprise, as were social scientists who have produced celebrated theories of social change with ostensibly substantial predictive power (for a detailed evaluation of these theories, see Lichbach [1995, esp. sec. 9-2]).

Another category encompasses the peoples of Eastern Europe. A vast number of impressionistic accounts suggest that the East Europeans themselves were stunned by the sudden collapse of their communist regimes. Some systematic evidence comes from a survey conducted four months after the breaching of the Berlin Wall. In March 1990, researchers from Germany's Allensbach Institute asked a broad sample of East Germans, "A year ago did you expect such a peaceful revolution?" Only 5% answered yes; 18% responded "yes, but not that fast"; and as many as 76% admitted to have been totally surprised.<sup>3</sup> These figures are all the more remarkable given what psychologists call the I-knew-it-would-happen fallacy—the human tendency to exaggerate foreknowledge (Fischhoff and Beyth 1975).<sup>4</sup>

With regard to the foregoing categories—outsiders and rank-and-file East Europeans—one might say, on the grounds that in a large society it is individually optimal to remain ignorant on social matters,<sup>5</sup> that people with little influence on the course of events, or with little stake in political trends, have no reason to invest in gathering information concerning the probability of revolution. Yet among those surprised were many individuals with everything to lose or gain from the fall of communism. They form our final two categories.

On the losing side, the surprised included the leaders of the incumbent communist regimes. Even without Soviet military assistance, the local communist parties had the means to suppress the growth of public opposition at an early stage. If they allowed public opposition to gather mo-

<sup>3</sup> East German Survey of the Institut für Demoskopie Allensbach (February 17–March 15, 1990), Survey 4195.

<sup>4</sup> In experiments conducted by psychologists, the I-knew-it-would-happen fallacy worsens over time; i.e., the greater the length of time that has elapsed since an event's occurrence, the larger the number of subjects that report they foresaw it. One would expect, therefore, the share of East Germans reporting that the explosion caught them by surprise to decrease with the passage of time. When asked in March 1991, a year after the first survey and 16 months after the fall of East German communism, "Two years ago did you expect such a peaceful revolution?" 7% answered yes and 33% said, "Yes, but not that fast." The share of respondents indicating that they were totally surprised was down to 54% (Allensbach Archives, Survey 5049). In March 1993, however, when a sample was asked, "Four years ago did you expect such a peaceful revolution?" the first two figures fell back to 4% and 23%, respectively, and the share of the totally surprised rose to 70% (Allensbach Archives, Survey 5078). It is too early, of course, to tell whether recollections of the East German uprising will follow the pattern observed in experimental studies.

<sup>5</sup> The logic of "rational ignorance" is developed by Downs (1957, chaps. 11–14).

mentum, a major reason is that they failed to recognize how quickly events could slip beyond their control. Remarkably, even after several of the East European regimes had fallen, Romania's Nicolae Ceaușescu went on a state visit to Iran, confident that his domestic support was secure and that he could easily block any attempted uprising.

On the winning side, the stunned included the East European dissidents. For reasons that will become clear, this is the most suggestive category, for some dissidents understood, as few others did, the immense vulnerability of their regimes. For example, in the late 1970s Czechoslovakia's Václav Havel ([1979] 1985) wrote a brilliant essay, "The Power of the Powerless," in which he argued that communism would some day fall like a house of cards, because it was sustained by widespread lying. Yet, when signs of the impending East European overturn began to multiply, as when Gorbachev hinted that the Brezhnev doctrine was dead, even an astute observer such as Havel dismissed the signs as insignificant. On several occasions before the events of late 1989, he counseled his readers to stop dreaming (see Kuran 1991, p. 9).

The East European revolutions were not, of course, the first major uprisings to catch the world by surprise. The French, Russian, and Iranian revolutions are only three of the successful revolts that stunned their leaders, participants, victims, and observers. One of the central points of Tocqueville's masterpiece, *The Old Régime and the French Revolution* ([1856] 1955), is that no one foresaw the fall of the French monarchy. Just weeks before the Russian Revolution of February 1917, Lenin suggested that Russia's great explosion lay in the distant future and that he himself would not live to see it. Even after the onset of the demonstrations that would bring down the Romanov dynasty, diplomats in St. Petersburg were informing their capitals that Russia remained stable. A few months before the Iranian Revolution of 1978–79, a U.S. intelligence report spoke of Iran as an island of stability in a sea of turbulence (Hoveyda 1980, pp. 16–17). For its part, the Iranian Communist Party was so off base in its forecasts that the Soviet government, which had been using the party as a source of information, sacked its entire leadership. Even the Ayatollah Khomeini was stunned by the events that propelled him to power. Although in public he was insisting that the shah's regime was on the brink of collapse, to his close associates he was confiding serious reservations until about two weeks before his triumphant return to Tehran (Heikal 1982, pp. 156–57).

#### SOME PUZZLES AND TWO CHALLENGES

Why might individuals with deep insight into a social system, or with privileged access to information about its undercurrents, fail to foresee

its impending explosion? What is it that can keep even the most astute and best informed members of a society unaware of imminent political changes of epochal significance? These questions are especially puzzling since now, in retrospect, various signs of the impending revolutions in France, Russia, Iran, and Eastern Europe are transparently obvious. With regard to the last case, we see the economic failures of the Soviet Bloc and the immense frustrations generated by the communist monopoly of political power. It is revealing that, just a few years after the revolution, there already exist multitudes of books and articles that present a vast array of reasons why East European communism was likely to, even had to, fall (see, e.g., Brown 1991; Frankland 1992; Ramet 1991). But if the signs of change are now so clear, why were they not noticed prior to late 1989? Why has our hindsight with respect to the collapse of East European communism proved vastly superior to our foresight?

Several years ago I sought to resolve these puzzles (Kuran 1991), drawing on a general theory of revolutionary surprise that I had developed before the fall of East European communism (Kuran 1989). Below I provide a sketch of that basic argument. The overview will set the stage for addressing two challenges.

The first challenge is to explore what our dismal predictive record implies for the possibility of predicting future revolutions. The answer is not at all obvious. The fact that our foresight has been weak in the past does not mean that all possible future revolutions are equally likely or equally unlikely. It does not imply that, looking ahead in time, we should attach equal probability to a political revolution in Austria as to one in Mexico, Romania, or Egypt. Nor, therefore, does it mean that our capacity for predicting sudden social change cannot be improved.

The article's second challenge is to identify techniques for improving our predictive powers. I shall argue that there are ways to identify the vulnerabilities of a political status quo and, hence, to narrow the range of uncertainty with regard to social change. At the very least, one can identify where future revolutionary surprises are unlikely to occur.

#### PREFERENCE FALSIFICATION AND THE CONCEALMENT OF REVOLUTIONARY POTENTIAL

What is the appropriate response to the observation that past revolutions have been unanticipated? Inferring that no one has yet developed an adequate model of social change in general, and of revolution in particular, one could try to develop a better model that would yield perfect or near-perfect predictions. Alternatively, one could incorporate the fact that revolutions tend to come as a surprise into the set of phenomena to be explained. In my view, the latter option is both more realistic and

more fruitful. Accordingly, the theory that I am about to summarize does not presuppose the potential availability of all pertinent information. Rather, it includes what people do and do not know among the factors to be illuminated. It seeks at once to produce knowledge and to identify the limits of knowledge.

The theory rests upon interdependencies among the decisions of political actors. Specifically, I recognize that on any given issue (whether trade liberalization, abortion laws, or the appropriateness of a political regime) a person has a private preference and a public preference. When the two differ, the individual is engaged in *preference falsification*. To take the simplest formulation, let an individual's private preference be known only to himself, and let his public preference represent what he chooses to reveal to others. The distributions of these two preferences over the population may be called, respectively, private opinion and public opinion.<sup>6</sup> In practice, a person's private preference on a given issue is influenced by the social processes that determine public opinion.<sup>7</sup> For our purposes here, however, we can take private preferences and their evolution as given. The simplification will enable us to focus on the formation and transformation of public preferences and public opinion. Even cursory descriptions of these processes will yield key insights into the puzzles outlined above.

One determinant of a person's public preference is his or her private preference. Preference falsification is costly to the falsifier, in that it entails a loss of personal autonomy and a sacrifice of personal integrity. Relative to a person who approves of the prevailing regime, then, a person who despises it is more likely, holding all else constant, to join an antigovernment rally. Another determinant of the person's private preference is the set of benefits and costs associated with alternative public-preference options. If the likely cost of joining the rally, and thus revealing a preference for political change, is a stint in jail or ostracism by one's peers, the prudent course of action may be to remain on the sidelines, even to cheer on the security forces to leave no doubt as to where one has chosen to stand. The external benefits and costs associated with a public-preference choice generally depend on the choices of others. If only a few people are demonstrating against the regime, the possible external cost of participation is likely to be much higher, and the expected benefit much lower, than if the streets are packed with demonstrators.

Typically, certain configurations of expected benefits and costs will keep an individual loyal to the incumbent regime; others will make him

<sup>6</sup> For fuller descriptions of these concepts, see Kuran (1995, chaps. 2–4).

<sup>7</sup> On these processes, see Kuran (1995, chaps. 10–14 and 17–18).

join the opposition. The configurations under which individuals are indifferent between the two options define their revolutionary thresholds. In any given society in which the regime's legitimacy is being challenged, individual members will differ in their revolutionary thresholds. Variations in private preferences are sufficient to generate a distribution of thresholds, as are variations in sensitivity to social pressure.

Under a very broad class of situations, the threshold distribution will generate more than one self-sustaining distribution of public preferences. In other words, public opinion will feature multiple equilibria. One or more of these equilibria may harbor revolutionary implications for the prevailing social order. If such equilibria exist, however, this will not necessarily be known because thresholds are not common knowledge.

A major implication is that, in any given society, cognitive, economic, and social processes may be making it ever easier to spark a revolutionary bandwagon without anyone sensing the potential for social change. The society may be on the verge of a massive explosion, therefore, with everyone continuing to believe—and indeed its members continuing to claim—that it is quite stable. At some point, a small, intrinsically insignificant event will suffice to activate a revolutionary bandwagon—although there is no guarantee that the event will actually occur. If the bandwagon begins rolling, it will catch everyone by surprise, including the very individuals whose actions put it in motion.

When a revolution occurs, long-repressed grievances burst into the open. Moreover, people who were relatively content with the old regime embrace the new one in order to avoid being stigmatized and persecuted as potential counterrevolutionaries. They pretend that their support for the old regime was never genuine, that it involved preference falsification motivated by self-preservation. In the process, they make the toppled regime appear even more vulnerable than it actually was. And they make it doubly easy for scholars to concoct plausible explanations for the observed events. So it is that the social sciences regularly produce multitudes of explanations for revolutions that no one had predicted. This paradox is rooted, I submit, in preference falsification. Before a revolution, preference falsification conceals the potential for a successful revolt. After the fact, it masks the factors that had been working against change.

#### THE LIMITS OF SOCIAL KNOWLEDGE

Preference falsification thus distorts and conceals information about political possibilities. What benefits could flow, one might wonder, from a theory that exposes obstacles to social understanding and forecasting? Why explore the limits of social knowledge when so much that is clearly

knowable remains uninvestigated? And is it not premature to accept defeat on the matter of explaining and predicting revolutions?

I do not share the apprehensions reflected in such questions, for identifying the limits of knowledge is not a declaration of failure. On the contrary, it is itself a contribution to the pool of useful knowledge. It is also, I would add, a necessary step toward charting a realistic scientific agenda. Darwin did not obstruct or retard science by developing a theory that limits our ability to predict the future course of biological evolution. Producing a quantum leap in biological knowledge, he catalyzed research that has actually improved our capacity to control the evolution of particular species. The goal of all science, not just biology, should be to explain the explicable, predict the predictable, and, equally important, separate the knowable from the unknowable.<sup>8</sup>

Objectors should scan the major social trends of our time and ask themselves whether any had been predicted accurately and whether, even with the benefit of hindsight, any are understood fully. Do we know exactly why the late 20th century has seen a rise in religious fundamentalism, an intensification of nationalism, and the rise of East Asia as an economic giant? Who foresaw, back in the 1950s, that by the 1970s South Korea would be an industrial exporter to reckon with, secularism would be on the defense throughout the world, and ethnic warfare would have escalated? Even a cursory evaluation of the past record of the social sciences will show that, at least on politicized matters involving multitudes of decision makers, neither perfect prediction nor full explanation is the norm. Not even economics, arguably the most disciplined and most advanced of the social sciences, has had much predictive success in domains involving collectively decided outcomes. Theories that explain how economies operate, like Adam Smith's ([1776] 1937) theory about how self-interested agents produce economic order and prosperity, do not furnish credible long-range predictions regarding the economic fortunes of nations.<sup>9</sup>

As a practical matter, then, the dichotomy between social theories that produce complete knowledge and ones that generate only limited knowledge is more apparent than real. The critical difference is that

<sup>8</sup> This methodological position conflicts in at least two respects with Milton Friedman's (1953) view that predictive success is the sole test of a theory. Where Friedman will settle for manifestly unrealistic "as if" assumptions for the sake of good predictions, I regard assumptions as critical to sound explanation. Second, where Friedman regards predictability as essential to the scientific enterprise, I consider it important that science recognize its own limits.

<sup>9</sup> For more on the disjunction between the explanatory successes and predictive failures of economics, see Hahn (1993). The dynamic processes that underlie this disjunction have been surveyed and explored by David (1993).

some theories disguise their limitations while others make them explicit.

The limitations of the outlined model of revolution stem from two distinct factors. The first consists of the model's nonlinearity, which is due to interdependencies among the public preferences that form public opinion. Specifically, the sensitivity of public opinion to changes in individual dispositions is variable rather than fixed. Nonlinearity allows huge variations in the consequences of a given perturbation. It permits the effects of the perturbation to be disproportionately large or disproportionately small. Within the present context, massive changes in private opinion may leave public opinion undisturbed, only to be followed by a tiny change that transforms public opinion radically through a bandwagon process. The trigger that activates a bandwagon effect may be events in other societies. In 1989, each successful challenge to communism lowered the risk of open dissent in countries still under communist rule. The consequent relaxation generated a domino effect, with a bandwagon in one country touching off even speedier bandwagons elsewhere.<sup>10</sup>

The second limit on the model's explanatory and predictive powers is the imperfect observability of the interdependencies among public preferences. We can never be certain, especially in advance, about the effects of a given switch to the opposition. Where preference falsification is rampant, then, we may fail, however well we understand the pervasiveness and implications of preference falsification, to notice that an incumbent equilibrium is about to vanish. Remember that Havel, who recognized that support for communism was mostly feigned and who understood the consequent vulnerability of the incumbent regimes, did not foresee the events of late 1989.

Let us be clear about the effects of nonlinearity and imperfect observability. Absent preference falsification, which is the source of imperfect observability, we would always be aware of approaching breaks in a society's political evolution. We would know, for instance, that an anti-government bandwagon would begin rolling if just 100 more people became disillusioned with the political status quo. And absent interdependencies among public preferences, the source of nonlinearity, small changes in individual dispositions would not produce explosive changes in public opinion. Whereas an easily observed change in an entire society's stake in the status quo might well precipitate a major shift in public opinion, an unnoticeable change in just a few people's feelings would generate at best a commensurately tiny shift. With both preference falsi-

<sup>10</sup> The domino effect has been analyzed by Kaempfer and Lowenberg (1992).

fication and interdependencies, effects can be unforeseen and disproportionate.

The fact that an outcome was unforeseen does not imply that it must remain a complete mystery. With the benefit of hindsight, many unanticipated events are understood reasonably well. We know much about the conditions and frustrations that propel modern fundamentalisms, even though their rise was scarcely predicted. We know why and how the East Europeans brought down their communist regimes, even though the breaching of the Berlin Wall amazed us all. The above-summarized theory explains, in a manner compatible with its own logic, why explanation is often easier than prediction: a shift in public opinion brings to the surface a plethora of information that is consistent with the shift just as it conceals information that is inconsistent with the shift.<sup>11</sup> The problem is compounded by the fact that data consistent with an event are noticed more readily than are inconsistent data (see Fischhoff and Beyth 1975; Nisbett and Ross 1980, esp. chaps. 5–7).

A model's ability to demonstrate why explanation is easier than prediction should not be taken lightly. In the social sciences the two concepts are often used interchangeably, as though a model that yields insights into the past must be equally good at predicting the future. Moreover, retrospective accounts seldom make clear what actors actually knew and what they could have known. Such accounts generally suggest that recorded events just had to happen, failing to explain why, if so, the events were not predicted.

As a case in point, many accounts of the East European revolutions propose that they were inevitable. These accounts are all very misleading. If the old communist order were still in place, would we not be advancing persuasive reasons for the permanence of communism? Who would have paid serious attention to signs of communism's vulnerability? The truth is, if no revolution had occurred, few observers of Eastern Europe would have considered the ongoing stability of Eastern Europe a puzzle.

For another revealing case, consider the recent history of Western scholarship on the Iranian state. Until the Iranian Revolution, almost all students of Iran subscribed to the view that the Iranian state was traditionally despotic and the Iranian nation submissive and fatalistic. Beginning in 1979, many started depicting Iran as a country whose history is marked by a strong society and a weak state. The Iranian clergy,

<sup>11</sup> There exist theories that predict better than they explain. For example, the Ptolemaic theory of the universe is quite successful at predicting the movement of planets, but by modern standards its explanation for these movements is very inadequate (see Kuhn 1957).



until 1979 portrayed as quietistic, was now said to control a revolutionary force capable of making and breaking regimes.<sup>12</sup> If Iran were still ruled by the Pahlavis, today's fashionable theories would probably be confined to the fringes of Iranian studies.

From the practical difficulties of social prediction some scholars infer that general social theories are useless. We should limit ourselves to case studies, they say, without seeking to achieve conceptual unity. In the context of revolutions, for example, repeated predictive failures have made some political scientists call for an end to the quest for a general theory of revolutions (see, e.g., Dunn 1989, pp. 2–3; Bunce 1991, esp. pp. 152–53). Such writers are to be applauded for drawing attention to our dismal record at predicting social upheavals. They are wrong, however, to dismiss the possibility of general insights into the revolutionary process. The problem lies not with theorizing *per se* but with the type of theorizing that has dominated the social sciences. What we need are theories that account explicitly for their own limitations and that distinguish between explanation and prediction. Theories that meet these two criteria are capable of accommodating the observations that are said to prove the futility of searching for conceptual unity and generality.

It will be instructive, before we move on, to contrast the notion of imperfectly observable nonlinearity with its polar opposite, observable linearity. Let  $a$  and  $b$  be observable variables, and consider the linear equation  $b = 3a + 2$ . This equation indicates that changes in  $a$ , however small or large, result in three times larger changes in  $b$ . As such, it cannot capture variations in the sensitivity of  $b$  to changes in  $a$ . Variations might be accommodated, of course, through “noise”—the statistician's euphemism for chance events, data imperfections, and just plain ignorance. The equation  $b = 3a + 2 + \epsilon$  indicates that the effect of a change in  $a$  may vary by  $\epsilon$ , the amount of noise. When  $\epsilon$ , whose determinants may be unknown, is 0.5, a jump in  $a$  will produce a change in  $b$  that is more than three times as large.

Readers will recognize the second equation as the type used in linear regressions. Does a regression accommodate, by virtue of its noise term, the possibility of surprise? It does. Yet it attributes surprises to unexplained noise rather than to a specific social process. Also, statistical noise is meant to accommodate small surprises, not enormous ones of the sort we faced in late 1989. Only a nonlinear model featuring imperfectly observable variables can capture the fact that revolutions often catch us unprepared.

<sup>12</sup> The point is developed by Sadowski (1993, p. 16), who offers several additional examples of changing scholarly fads in response to unanticipated events.

#### TOWARD IMPROVED EXPLANATION AND PREDICTION

Because preference falsification afflicts every society, albeit in varying forms and degrees, we will be surprised by future revolutions. Moreover, we are likely to face, after each unanticipated explosion, an abundance of information pointing to the high likelihood, if not the inevitability, of the observed course of events. Interpreters will notice the warning signals of the impending revolution, but not the contradictory and competing signals. And such retrospective distortions will be compounded by citizens' postrevolutionary incentives to hide their past and present sympathies for the toppled order.

To identify limits of knowledge is not to say, of course, that we are doomed to total ignorance about the past. Nor is it to propose that unfolding events must always amaze us. The limits of knowledge about revolutions stem not from the unobservability of the interdependencies among public preferences, but rather, the imperfect observability of those interdependencies. The signs of preference falsification are seldom fully hidden, and the pressures that promote it are often partially identifiable. Depending on the context, for instance, much relevant information may lie in the results of anonymous polls. Sometimes, helpful information may be drawn from such sources as memoirs, diaries, confidential letters, deathbed confessions, and secret archives. One can often distinguish, therefore, between genuine social consensus and suppressed controversy and between honest devotion to a regime and fear-driven obedience.

Private opinion is generally harder to measure and interpret, it might be said, in countries where democratic freedoms are limited or fragile than in countries with strong democratic traditions. In the former the very forces that discourage truthful expression also inhibit the collection and dissemination of opinion data. As a case in point, the communist regimes of Eastern Europe regulated opinion research. They used various means, moreover, to mislead the world about the nature of private opinion. Nevertheless, we knew all along about the existence of widespread hidden opposition to communism. Most of us realized that the East European regimes enjoyed less genuine acceptance than, say, the French government did in France. What we could not know are the exact characteristics of East European private opinion, to say nothing of the precise distribution of individual revolutionary thresholds.

Not that preference falsification is a problem for social analysis only with respect to undemocratic countries. Even where the right to express offensive views enjoys legal protection, there exist sensitive issues on which people think twice before venturing an idea in public. In the United States, race relations represents just such an issue. We know this

through data ranging from scientific polls to everyday observation. Over the past quarter century, scores of scientific surveys have found that the racial admission and employment policies instituted under the rubric of affirmative action are overwhelmingly unpopular (Sniderman and Piazza 1993, esp. chaps. 4–5). Yet overt opposition to racial affirmative action is rare (for a detailed analysis of preference falsification in American race relations, see Kuran [1995, chaps. 9, 14]).

Where open and honest conversation is blocked, there will generally exist signs of hidden opposition to positions that enjoy vast public support. One can identify, therefore, the existence of a potential for social explosion. So, notwithstanding its acknowledged limitations, the theory I have outlined does enhance both explanation and prediction. It can improve our readings of history and alert us to future possibilities.

However democratic or undemocratic, all polities feature openly contested issues on which one can express a wide range of views without enduring significant penalties. On such issues public opinion will not make unforeseen jumps, except in response to a major shock that changes many minds simultaneously, as when an earthquake jolts millions into reconsidering the importance of strict building codes. Also, past changes in public opinion can be explained without worrying much about poorly observed interdependencies. It is on socially sensitive issues that preference falsification may sharply limit our predictive and explanatory capabilities. Yet even on such issues, I repeat, we are never totally in the dark. With respect to the past, we can understand the persistence and the consequences of repressive conditions. With respect to the future, we know where to look for possibilities of sudden change.

The task of identifying instances of widespread preference falsification will generally involve the consideration of data difficult to document and interpret. Before 1989, a scholar trying to demonstrate the pervasiveness of East European preference falsification might have invoked (1) opinion surveys conducted by Western organizations on East European travelers, (2) the claims of dissidents like Havel, Adam Michnik, Andrey Sakharov, and Aleksandr Solzhenitsyn, and (3) the observations of informed outside observers. All such data could have been, and were, discounted as biased. They could not be supplemented with credible opinion surveys, however, because the governments in power refused to grant the necessary permissions. Not that polling data were nonexistent. The secret services of the communist regimes regularly conducted surveys to keep themselves informed about the true thoughts and feelings of the citizenry. Before 1989, however, these data were kept from the public. Now we understand, and then we had reason to suspect, that communist governments were secretive precisely because they recognized their private un-

popularity. A regime that enjoys genuine legitimacy has no reason to keep its opinion surveys classified. Nor does it need to prohibit independent polling.

It may appear unscientific to assert, when systematic polling data is unavailable, that a regime, institution, policy, or political agenda is privately unpopular. But the scientific ethos demands only that we gather the best data available and interpret our evidence in the light of sound theory. It does not require us to ignore problems on which data are relatively scarce or imperfect. The fact that the obtainable evidence on Eastern Europe was less precise than, say, data on the average class size in West European schools was no reason to ignore the realities of communism. In any case, for reasons already noted, the unavailability of good opinion data was itself an important sign of preference falsification. Just as it was significant to Sherlock Holmes that the dog did not bark, it was politically significant that, unlike West Germany, East Germany prohibited independent polling.

In contexts where preference falsification is rampant, then, we often have little choice but to employ whatever data can be found, even very inexact data. Whether interpreting the past or exploring future possibilities, we may have to pay attention to the scattered perceptions of observers who seem well informed about happenings behind the public stage. Impressionistic accounts do not provide the precise number of people who despise the status quo, but at least they point to the existence of widespread discontent.<sup>13</sup> After repression eases or disappears, of course, the relevant information is likely to improve. Since 1989 we have gained access to many secret surveys conducted, over the years, for the benefit of Eastern Europe's communist regimes.

Social predictions based on perceptions of preference falsification may suffer, it deserves mention, from a problem absent from historical explanation. Predictions interact with the phenomena they predict. However accurate its reading of social undercurrents, a report that society is about to erupt may become self-fulfilling; or, by provoking the government to take precautionary measures, it may become self-negating. Such effects are not specific, however, to contexts where people are afraid to express themselves truthfully. Any social observation may affect what is being observed.<sup>14</sup> An economist who predicts a rise in unemployment may contribute to a recession that would not have occurred had he kept quiet. And reports of a candidate's invincibility may scare off her most qualified opponents, thus compounding her advantage.

<sup>13</sup> A similar point is developed by Bermeo (1992, pp. 184–87).

<sup>14</sup> For an illuminating discussion, see Carr (1962, pp. 90–91).

## MEASURING PREFERENCE FALSIFICATION

To suggest that the concept of preference falsification can improve our predictive and explanatory capabilities is not to say, then, that it can be incorporated into social analysis risklessly. A further problem is that we lack a sufficiently developed set of scientific techniques for identifying and quantifying preference falsification. The latter problem does not necessarily point, I hasten to note, to a flaw in the foregoing argument. Techniques for forming new databases rarely get developed until new theories establish their usefulness (Lakatos 1978). Methods for quantifying temperature were devised only after physicists crafted theories featuring temperature scales. Techniques for measuring the velocity of money emerged only after the concept became commonplace in economic texts.

When a theory precedes the measurement techniques needed to verify or use it, it will generally be harder to test in historical contexts than in contemporary ones. Data on the velocity of coinage in 14th-century Anatolia are not as reliable as information on the modern velocity of the Japanese yen. This is reason for caution in Ottoman economic research, however, not a cause for spurning the concept of monetary velocity. Similarly, one can acknowledge the difficulties of measuring discrepancies between private and public opinion in the past, especially the distant past, without declaring the concept of preference falsification useless. The fact that no scientific opinion surveys were conducted during the French Revolution does not mean that it is pointless to introduce the concept of public opinion into research on 18th-century France.

There have been attempts, in fact, to identify cases of preference falsification in the distant past. For example, medievalists have developed a technique for determining concealed messages in old philosophical texts. It is based on a simple principle: when an experienced writer expresses a view at odds with public opinion, one has reason to believe that he is expressing himself truthfully; but when he conveys a view in line with public opinion, one cannot rule out the possibility that he was trying to avoid punishment, especially if the view contradicts what he wrote elsewhere. Medieval philosophers wrote at a time when challenges to widely held beliefs often brought swift retribution. Under the circumstances, they made it a practice to present their most original and potentially most controversial thoughts "between the lines" for the exclusive benefit of other independent thinkers. As Leo Strauss (1952) and others have documented, in the depths of a treatise an able writer would surreptitiously contradict an orthodox tenet that he had defended in many conspicuous passages, with an eye toward exposing his dissent to cultivated readers likely to be sympathetic, while simultaneously hiding it from unsophisticated readers likely to be offended. Careful readings of

the works of Farabi (Alpharabius), Maimonides, Ibn Khaldun, Hobbes, Spinoza, and other towering philosophers suggest that they tended to express their heretical thoughts in relatively inconspicuous passages, and in deliberately ambiguous terms, probably to escape persecution.

Reading between the lines is not an infallible technique. But to reject it for this reason alone would be like denying emergency aid to a wounded sailor at sea on the grounds that he could get better treatment from a fully equipped hospital on shore. The value of any datum is contingent on what else is available. In any case, to presume that past writers had neither incentives nor a capacity to hide their true thoughts would contradict certain basic facts of human nature. Moreover, it could lead to serious historical misinterpretations.

The point remains that in historical contexts we rarely have access to ideal data. If future historians are to have better data on our own age, we must make systematic efforts to collect data that distinguish between private and public opinion. In the meantime, such data may help us improve our predictive capacities with regard to social stability and instability. In particular, they may contribute to the identification of societies that are prime candidates for political upheaval.

The essential tasks consist of qualitative field research and quantitative surveys. Field research would be performed by scholars trained in the anthropological techniques of "thick description"—establishing rapport with a community, selecting informants, keeping a diary, and so on. Living for awhile in the community under investigation, they would try to win the community's trust in order to gain exposure to the perceptions, ideas, resentments, aspirations, and ambitions that its members tend to keep private. The research would thus capture differences between the community's life on stage and its life off stage. The political scientist James Scott (1985, 1990) has shown how this can be done through his own fieldwork in Malaysia. He has documented that poor peasants deliberately and routinely mislead their landlords and government officials about their knowledge, aspirations, and resentments.

As for quantitative surveys, certain techniques that contribute to predicting and explaining revolutions have already been developed. I shall discuss two, beginning with one of the many tests developed by the Allensbach Institute under the leadership of its founding director, Elisabeth Noelle-Neumann. A representative Allensbach technique is the "two-stage train test" (Noelle-Neumann [1980] 1984, pp. 16–22). In 1972 this test was used to measure the incentives to take public positions for and against *Ostpolitik*, Chancellor Willy Brandt's attempt to come to terms with communist Eastern Europe. In the first stage of the test, a sample of individuals were presented with a sketch showing two people in conversation, one making a statement favorable to *Ostpolitik* and

the other an unfavorable statement. Each respondent was then asked individually to state his or her own opinion. In the second stage of the test, both those in favor of *Ostpolitik* and those against it were invited to imagine being at the start of a five-hour train ride. Half of each group heard that they would have a pro-Brandt compartment mate, the other half that their compartment mate would be anti-Brandt. Each interviewee was then asked: "Would you like to enter into a conversation with this person so as to get to know his or her point of view more closely, or wouldn't you think it worth your while?" The goal, of course, was to measure the perceived social pressures. People do not like to think of themselves as cowards, so if asked about their fears directly they may seek to hide them. The train test gets at the fears indirectly, and it allows the quantification of opposing pressures.

Of the survey sample, 50% of those favorable to Brandt's policy, as against 35% of those unfavorable, indicated a preference for entering into a conversation. Correspondingly, "would not think it worthwhile" was the answer given by 42% of the favorable group, as against 56% of the unfavorable group. At the time of the survey, the Brandt camp appeared larger than the anti-Brandt camp. The survey results suggest that, to a degree, the perception reflected the fact that pro-*Ostpolitik* citizens were relatively more willing to state their positions publicly. Evidently, preference falsification was benefiting the cause of *Ostpolitik* by tilting public discourse, and thus public opinion, in its favor.

Let me turn now to another survey technique, developed, like the train test, as a means of predicting an election outcome. In the heat of an election campaign social pressures favoring one side or the other may give misleading signals as to the outcome of the pending election. The induced preference falsification may bias preelection polls, especially if polltakers are suspected of dangerous partisan sympathies. As a case in point, a Washington Post-ABC News poll completed 10 days before the Nicaraguan election of 1990 gave the Sandinista presidential candidate, Daniel Ortega, a 16-percentage point lead over the candidate of the Union Nacional Opositora (UNO), Violeta Chamorro. Other polls gave Ortega an even wider lead. Chamorro would go on to win the election by 14 points. Yet, taking the projections at face value, many news organizations had interviewed Sandinista leaders right before the election to discuss how they planned to exploit their imminent victory. The only polls that turned out to be in the right ballpark were ones conducted by organizations linked, in fact or in the popular imagination, to UNO.<sup>15</sup> Foreign news organizations had dismissed the latter polls as partisan.

<sup>15</sup> For evidence, see Miller (1991) and also Schwartz (1992, esp. chaps. 4, 7)

Consequently, they were stunned by Chamorro's victory, much as they had been stunned just a few months earlier by the fall of East European communism (Uhlir 1990; Ornstein 1990).

An ingenious experiment run by Katherine Bischooping and Howard Schuman (1992) points to the source of confusion. A few weeks before the election, Bischooping and Schuman conducted 300 interviews in and around Managua, all administered identically except for the type of pen used to record responses. In one-third of the interviews the interviewer used a pen featuring the red and black colors of the Sandinista Party and the inscription "DANIEL PRESIDENTE." In another one-third the interviewer used a pen featuring the blue and white colors of the opposition and the inscription "UNO." And in the remaining one-third the interviewer used a neutrally colored pen with no lettering. Interviewers did not draw attention to their pens or make claims about their own political sympathies. Yet, the results show clearly that the pens influenced the respondents. When the interviewer held a Sandinista pen, the respondents voiced support for Ortega by a 26-point margin. Ortega also came out ahead in the neutral-pen condition, by 20 points. However, when the interviewer held a UNO pen, Chamorro was the winner by 12 points.

The UNO-pen condition thus came close to predicting the election outcome, whereas the Sandinista-pen condition replicated the highly inaccurate preelection polls. Remarkably, the neutral-pen condition generated a result similar to the Sandinista-pen condition. Bischooping and Schuman suggest that after a decade of Sandinista repression a pollster lacking partisan identification would have been perceived as a Sandinista activist. Insofar as their intuition is correct, we have an explanation for why the Washington Post-ABC News poll was so far off the mark. Precisely because it was conducted by interviewers striving for a neutral image, respondents sympathetic to UNO considered it prudent to keep their private preferences concealed. Evidently, only interviewers with apparent UNO connections gave UNO sympathizers the courage to reveal the truth. Sandinista sympathizers needed no such assurance, for Chamorro was stressing her readiness to tolerate dissent. And in any case, widespread preference falsification on the part of UNO sympathizers had lowered her perceived chances of winning.

Like the Allensbach train test, the pen experiment identifies fears and sensitivities. It thus alerts one to possible incongruities between public and private opinion. Since elections by secret ballot measure private opinion, polls undertaken to predict electoral outcomes will yield misleading forecasts unless the respondents feel comfortable expressing themselves honestly. Of course, interpreting a poll designed to overcome preference falsification is anything but a mechanical matter. Drawing sound inferences from the pen experiment required an understanding of



Nicaraguan political realities. An analyst ignorant of Nicaraguan politics may well have designated the source of fear as UNO, for it was the UNO pen that registered a dramatically different outcome from the neutral pen. One needs to know the history of Sandinista rule—the use of informants, the persecution of opponents, the biases of the media—to see that a pollster trying to appear neutral would probably be perceived as pro-Sandinista. So, while clever experiments may provide valuable social insights, they are not self-explanatory. Analysts in agreement that an experiment points to widespread preference falsification might disagree on the character of the falsification.

For all their ambiguities, experiments like those just discussed have uses also in nonelectoral contexts. They can be brought into broader studies of social stability and political evolution. They can be used to identify possible political upheavals and to explain ones that have already occurred. Oddly, contemporary students of political revolution have taken little advantage of developments in the field of opinion studies. Yet, as I have endeavored to show here and elsewhere (Kuran 1989, 1991), the dynamics of private and public opinion offer many clues into the process of social change.

Measurement techniques like the ones I have illustrated may be criticized on the grounds that they yield imprecise readings and that they are not standardized. Let me address each point in turn.

It is true that pen experiments and train tests provide inexact readings of private opinion, preference falsification, or perceived social pressures. But I am not suggesting otherwise. Because I recognize that the determinants of political outcomes are imperfectly observable, I am arguing merely that techniques exist to provide rough readings of the political climate and crude estimates of private opinion. Techniques that yield more exact measurements may yet be developed. However, barring the invention of an instrument for reading the individual mind, they will never attain microscopic precision. In any case, not every domain of analysis requires the same precision. Just as the appropriate unit for measuring the distance between two stars is not the micron, we do not need to know private opinion exactly to uncover the existence of a latent revolutionary bandwagon effect. If a variant of the pen experiment had been run in East Germany in 1988, and the nature of the interviewer's pen had been found to make a huge difference, the finding would certainly have been informative. We would still not have known that the Berlin Wall would be breached in a year's time. But at least we would have obtained controlled evidence against the argument that the Soviet Bloc was incapable of self-driven change.

It is also true that I have not offered standardized procedures for measuring hidden variables—procedures that can be applied more or less

mechanically to every possible situation. The train test was developed for a country where people commonly ride the train. It would not be as meaningful in a place where few people even lay eyes on trains, except in movies. Perhaps standardized tests will someday be developed and put into regular use. They will still need to be interpreted, however, in the light of their political and social contexts. Inflation figures can be compared meaningfully without attention to background conditions. If the Mexican inflation rate is double the Turkish rate, we know something specific about the respective price movements. But train tests conducted in Mexico and Turkey would be meaningful only in the light of information about social and political conditions in the two countries. Identical results might signify widespread political fear in one country, customary politeness in the other.

#### REFUTABILITY OF THE FOREGOING PROPOSITIONS

It is time to pull together the threads of the foregoing argument. I have advanced three propositions.

PROPOSITION 1.—*The ubiquity of preference falsification makes more revolutionary surprises inevitable.*

PROPOSITION 2.—*Unanticipated regime changes will tend to occur in politically repressive countries—ones whose regimes enjoy little genuine legitimacy and are sustained by general fear. Politically repressive countries can be identified through techniques for detecting instances of widespread preference falsification.*

PROPOSITION 3.—*Obstacles to predicting revolutions preclude neither the explanation of particular revolutionary surprises nor the development of elaborate retrospective accounts of mass uprisings.*

Each of these propositions makes a refutable claim. The first proposition can be debunked simply by constructing a theory that predicts future revolutions accurately. Here are some examples of predictions that could cast doubt on the claim:

1. In Egypt, Hosni Mubarak's regime will be overthrown between 1995 and 1997.
2. Brazil and Venezuela will experience mass uprisings around the year 2000, whereas Mexico, Colombia, and Ecuador will remain stable.
3. The communist monopoly of power in China will end abruptly during the next five years, but North Korea's communist regime will manage to hold onto power for another four decades.

If successful, such a string of predictions would suggest that preference

falsification is not, contrary to what I maintain, an obstacle to foreseeing revolutions.

The second of my three propositions would be discredited if unanticipated political revolutions began to occur primarily in countries where political freedoms are widely respected, rather than in politically repressive countries. Since preference falsification with respect to the incumbent regime is a hallmark of politically repressed countries, we would have to conclude that preference falsification is not, after all, the main obstacle to prediction. Finally, my third proposition would be refuted if future historians began having difficulty explaining unanticipated revolutions.

#### STRUCTURE, CHOICE, AND THE PERMANENCE OF OBSTACLES TO PERFECT PREDICTION

Although I have argued that our dismal record at predicting revolutions may be improved marginally, some scholars maintain that, armed with the right model, we can do very much better. For instance, Jack Goldstone (1993) holds that, despite the impediments to observing private opinion, it is possible to identify the "objective conditions" for revolution. He bases the claim on a "demographic-structural" model that he has used to explain various political instabilities in the early modern world, including revolutions and rebellions in England, France, Turkey, and China (Goldstone 1991). If he is right and I am wrong, his demographic-structural model (or some other model) will begin yielding precise predictions. The predictions would have to be genuine, of course; that is, they would have to be prospective rather than retrospective. And they would have to be accompanied by successful predictions of political stability. Future decades should provide sufficiently many new cases for impartial observers to evaluate whether his optimism about predicting regime changes is warranted. The empirical validation of Goldstone's proposition would jeopardize my own first proposition.

Another thinker who holds that future revolutions can be predicted is Randall Collins (in this issue), who bases his claim on a "geopolitical theory" that he used, in 1980, to predict that the Soviet Union would disintegrate "within 30–50 years." Though Collins himself was "surprised that it happened so soon" (which is consistent with my thesis and inconsistent with his), he maintains that his theory can continue to provide accurate forecasts within "units of about 30–50 years." This is a testable hypothesis. If Collins ventures to propose what macropolitical changes will occur between 1995 and 2045, future scholars may be invited to judge, perhaps in this journal's sesquicentennial issue, the success of his projections. On the basis of arguments I have advanced here and elsewhere, I would expect most of the projections to turn out false. Preference falsification

is one constraint on predictability in the medium run. Another is that the private preferences that feed into public opinion are subject to change over time, partly in reaction to shifts in public opinion itself.

It is notable that Collins allows for a range of unpredictability of up to half a century. Yet such a range does not flow from the logic of geopolitical theory. One has to step out of geopolitics and appeal ultimately to the pervasiveness of preference falsification to explain why short-term geopolitical predictions are bound to be highly unreliable. This, I would argue, is a serious shortcoming. The objective of science is not only to predict but also to explain. A model that does not explain its own limitations is missing something important.

As I have suggested on other occasions (Kuran 1989, 1991), structuralist theories provide valuable insights into processes of political change. Structures do indeed affect political possibilities, partly by helping to shape people's incentives to revolt against an established order. Nonetheless, structuralist theories will remain incomplete even as explanatory devices until they are linked to a model that can account for poorly observed interdependencies among individual political choices.

Structuralists will answer, perhaps, that political structures do not just help shape people's incentives to revolt; they fully determine how individuals behave. Putting aside the question of whether there exists such a thing as free will, I would point out that the structures that constrain individual choices do not have lives of their own. They are the possibly unintended products of earlier individual choices. Just as market outcomes result from interactions between supply and demand, political outcomes result from interactions between social opportunities and individual dispositions. A critical feature of the latter interactions is that minor structural variations can make all the difference between continued political calm and an explosive growth in political opposition to the status quo.

If we could know exactly how the members of a given society would react to any given structural shock, we could foretell the society's political future. In practice, however, such knowledge is unobtainable. And an important part of the reason is precisely that individuals often conceal their political desires and motivations. Might the structuralists explain their predictive failures without appealing to considerations of individual choice? A possible attempt at rescuing the approach might be to suggest that, when many structural changes occur simultaneously, the consequences will be hard to predict. But one would have to go outside the structuralist framework to explain why multiple shocks hinder the task of prediction. And sooner or later one will have to come to terms with the impossibility of knowing with certainty how individuals will perceive the changes and how they will react to them.

I do not mean to imply that it is only the structuralist approach that exaggerates what is foreseeable. Certain arguments that share this article's commitment to methodological individualism make the very same mistake, albeit for different reasons. An example is the theory of revolutions developed by Susanne Lohmann (1994), which attributes revolutions to "informational cascades." The gist of her argument is that a revolutionary bandwagon effect develops as the potential opponents of the incumbent regime draw on their observations of others' political choices to make inferences about social conditions and political possibilities. Lohmann acknowledges that individuals are cognitively limited. The "private information" on which people base their political choices consists, she says, of a limited number of variables, including "the size of the protest movement and its changes over time." Yet, her entire argument rests on the empirically untenable notion that individuals can make fine distinctions among the characteristics of demonstrators. As she would have it, an East German citizen was able to look out his window at a crowd of antiregime demonstrators and estimate rather accurately its shares of "moderates" and "extremists."

Even more troubling is that Lohmann draws from her model—whose strength lies in the insights it provides into the oscillations that may occur in opposition size—the inference that demonstrations occur in "predictable ways." I doubt that anyone, not even a specialist in political revolutions, can look at a demonstration in the making and state with certainty where it will lead. Ultimately, of course, this is an empirical matter. Insofar as Lohmann's model has predictive power, it will doubtless become a prized instrument of political forecasting. My own expectation is that, like all the other theories of revolution considered in this article, including my own, Lohmann's model will derive more of its value from explaining revolutions than from predicting them.

## REFERENCES

- Ash, Timothy Garton. 1990. *The Magic Lantern: The Revolution of '89 Witnessed in Warsaw, Budapest, Berlin, and Prague*. New York: Random House.
- Bermeo, Nancy. 1992. "Surprise, Surprise. Lessons from 1989 and 1991." Pp. 178–200 in *Liberalization and Democratization: Change in the Soviet Union and Eastern Europe*, edited by Nancy Bermeo. Baltimore: Johns Hopkins University Press.
- Bischooping, Katherine, and Howard Schuman. 1992. "Pens and Polls in Nicaragua: An Analysis of the 1990 Preelection Surveys." *American Journal of Political Science* 36 (May): 331–50.
- Brown, J. F. 1991. *Surge to Freedom: The End of Communist Rule in Eastern Europe*. Durham, N.C.: Duke University Press.
- Bunce, Valerie. 1991. "Democracy, Stalinism and the Management of Uncertainty." Pp. 138–64 in *Democracy and Political Transformation: Theories and East-Central*

- European Realities*, edited by Gyorgy Szoboszlai. Budapest: Hungarian Political Science Association.
- Carr, Edward Hallett. 1962. *What Is History?* New York: Knopf.
- Collins, Randall. 1995. "Prediction in Macrosociology: The Case of the Soviet Collapse." *American Journal of Sociology*, in this issue.
- David, Paul A. 1993. "Path-Dependence and Predictability in Dynamic Systems with Local Network Externalities. A Paradigm for Historical Economics." Pp. 208–31 in *Technology and the Wealth of Nations: The Dynamics of Constructed Advantage*, edited by Dominique Foray and Christopher Freeman. London: Pinter.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Dunn, John. 1989. *Modern Revolutions: An Introduction to the Analysis of a Political Phenomenon*, 2d ed. New York: Cambridge University Press.
- Fischhoff, Baruch, and Ruth Beyth. 1975. "I Knew It Would Happen'—Remembered Probabilities of Once-Future Things." *Organizational Behavior and Human Performance* 13 (February): 1–16.
- Frankland, Mark. 1992. *The Patriots' Revolution: How Eastern Europe Toppled Communism and Won Its Freedom*. Chicago: Ivan R. Dee.
- Friedman, Milton. 1953. "The Methodology of Positive Economics." Pp. 3–43 in *Essays in Positive Economics*, by Milton Friedman. Chicago: University of Chicago Press.
- Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- . 1993. "Predicting Revolutions: Why We Could (and Should) Have Foreseen the Revolutions of 1989–1991 in the U.S.S.R. and Eastern Europe." *Contention* 2 (Winter): 127–52.
- Hahn, Frank. 1993. "Predicting the Economy." Pp. 77–95 in *Predicting the Future*, edited by Leo Howe and Alan Wain. Cambridge: Cambridge University Press.
- Havel, Václav. (1979) 1985. "The Power of the Powerless." Pp. 1–96 in *The Power of the Powerless: Citizens against the State in Central-Eastern Europe*, edited by John Keane and translated by Paul Wilson. Armonk, N.Y.: M. E. Sharpe.
- Heikal, Mohamed. 1982. *Iran: The Untold Story*. New York: Pantheon.
- Hoveyda, Fereydoun. 1980. *The Fall of the Shah*, translated by Roger Liddell. New York: Wyndham Books.
- Kaempfer, William H., and Anton D. Lowenberg. 1992. "Using Threshold Models to Explain International Relations." *Public Choice* 73 (June): 419–43.
- Kuhn, Thomas S. 1957. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge, Mass.: Harvard University Press.
- Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice* 61 (April): 41–74.
- . 1991. "Now out of Never: The Element of Surprise in the East European Revolutions of 1989." *World Politics* 44 (October): 7–48.
- . 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, Mass.: Harvard University Press, in press.
- Lakatos, Imre. 1978. "Falsification and the Methodology of Scientific Research Programmes." Pp. 8–101 in *The Methodology of Scientific Research Programmes*, by Imre Lakatos. Cambridge: Cambridge University Press.
- Lichbach, Mark I. 1995. *The Rebel's Dilemma*. Ann Arbor: University of Michigan Press.
- Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91." *World Politics* 47 (October): 42–101.

- Miller, Peter. 1991. "Which Side Are You On? The 1990 Nicaraguan Poll Debacle." *Public Opinion Quarterly* 55 (Summer): 281-302.
- Nisbett, Richard, and Lee Ross. 1980. *Human Inference: Strategies and Shortcomings of Human Judgment*. Englewood Cliffs, N.J.: Prentice Hall.
- Noelle-Neumann, Elisabeth. (1980) 1984. *The Spiral of Silence—Our Social Skin*. Chicago: University of Chicago Press.
- Ornstein, Norman. 1990. "Why Polls Flopped in Nicaragua." *New York Times* (March 7), p. A25.
- Ramet, Sabrina R. 1991. *Social Currents in Eastern Europe: The Sources and Meaning of the Great Transformation*. Durham, N.C.: Duke University Press.
- Sadowski, Yahya. 1993. "The New Orientalism and the Democracy Debate." *Middle East Report* 23 (July-August): 14-21, 40.
- Schwartz, Stephen. 1992. *A Strange Silence: The Emergence of Democracy in Nicaragua*. San Francisco: Institute for Contemporary Studies.
- Scott, James C. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven, Conn.: Yale University Press.
- . 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven, Conn.: Yale University Press.
- Smith, Adam. (1776) 1937. *The Wealth of Nations*. New York: Modern Library.
- Sniderman, Paul M., and Thomas Piazza. 1993. *The Scar of Race*. Cambridge, Mass.: Harvard University Press.
- Strauss, Leo. 1952. *Persecution and the Art of Writing*. Glencoe, Ill.: Free Press.
- Tismaneanu, Vladimir. 1992. *Reinventing Politics: Eastern Europe from Stalin to Havel*. New York: Free Press.
- Tocqueville, Alexis de. (1856) 1955. *The Old Régime and the French Revolution*, translated by Stuart Gilbert. Garden City, N.Y.: Doubleday.
- Uhlir, Mark A. 1990. "Nicaraguan Opposition Routs Sandinistas." *New York Times* (27 February), pp. A1, 12.

# Prediction in Macrosociology: The Case of the Soviet Collapse<sup>1</sup>

Randall Collins

*University of California, Riverside*

Macrohistorical predictions are possible, provided both a theory and empirical information are available. This article discusses sociological prediction, using geopolitical theory's successful prediction of the breakup of the Soviet Union as an example. Other predictions and explanations of the Soviet breakup generally lack theoretical validity. The precision of macropolitical predictions is limited to a range of decades, but state breakdown and revolution occur in a much narrower time period. The transfer of power happens in a mass mobilization lasting a few days, thus creating the illusion that spontaneous popular will causes macropolitical change and masking the structural shifts that make the change possible.

## GEOPOLITICAL THEORY AND THE COLLAPSE OF THE SOVIET UNION

### A Personal History of a Theory-Based Prediction

In 1978 I published a theory (Collins 1978) explaining changes in the territorial power of states. Extending conflict theory, I decided to take seriously Max Weber's definition of the state as the monopolization of legitimate force upon a territory. Turning this definition into an explanatory theory meant treating everything in it as a variable; the result was a theory of the conditions that determine geopolitical rises and falls in territorial power, together with the consequences that flow from these changes in power. A corollary of the theory is that the legitimacy of rulers varies with the external power-prestige of their state; at the extreme, this corollary entailed explaining revolution as the loss of legitimacy and of control over the means of coercion. The geopolitical theory thus meshed with Skocpol's (1979) state resource breakdown theory of revolution, published about the same time; the convergence between the two theories seemed to me additional evidence that the model was on the right track.

<sup>1</sup> This article was originally presented at the annual meeting of the American Sociological Association, Miami, August 1993. I am indebted to David Waller for advice in the preparation of this paper. Direct correspondence to Randall Collins, Department of Sociology, University of California, Riverside, California 92521-0419.



In 1980, a presidential election year, Ronald Reagan's major campaign issue was the so-called window of vulnerability: the claim that the United States had fallen dangerously behind the Soviet Union in nuclear armaments and needed a massive arms buildup to catch up. The early 1980s was the height of the period of nuclear terror, when the anti-nuclear arms movement mobilized under the image of "five minutes to midnight." I decided to apply my geopolitical theory to see what it predicted about our current situation. I honestly had no preconception of what the results might be.

The geopolitical theory comprised five principles of causal processes, interconnected by a cumulative dynamic. To my surprise, all five of the major principles in the theory indicated that the Soviet Union had passed the peak of its power and predicted that it would decline. The result was not symmetrical; most of these principles predicted that the power of the United States would remain relatively stable. Only one principle of the five held out the possibility that the United States would also decline, since nuclear war fitted within one of the more general categories of events by which state power is destroyed. My optimistic assessment was that the other four principles would take effect before the fifth and that the Soviet Union would disintegrate before a nuclear war could occur. The policy implication was that the nuclear arms race could be safely scaled back without undermining the relative power position of the United States.

In the spring of 1980, I presented this analysis at several places, including Yale and Columbia. The response was uniformly negative. Russia specialists, who attended some of these talks, were usually conservative émigrés, whose dominant feelings were hatred and fear of Soviet power. Their image was of a terrifyingly powerful Soviet Union, which must be combated by an equally powerful United States. This stance is not surprising from the point of view of Simmelian conflict theory, that an external threat brings ideological polarization and a cycle of escalation and counterescalation. The response of the liberals was a little more surprising. Some members of the nuclear disarmament movement reacted with hostility; during one talk, an activist accused me of sounding "just like the Joint Chiefs of Staff," apparently expressing the feeling that disarmament had to be justified as a moral crusade, not as an application of realpolitik. Perhaps more fundamental was the liberal position that the world was facing mutual assured destruction (MAD), based on the premise that the United States and Soviet Union were equally powerful and equally in need of deescalation.

Eventually I published the paper under the title "The Future Decline of the Russian Empire" in a volume of my collected essays, *Weberian Sociological Theory* (1986). There the prediction rested on the shelf. At

any rate, I was not surprised when the strains of the Afghanistan war brought the disgrace of the war faction in the Soviet Union, and their replacement by a reform movement around Gorbachev, nor when reform turned out to be a slippery slope toward disintegration of the empire.

The incident raises a number of general issues. To what extent is sociological prediction possible? How can we differentiate valid prediction from lucky guesses and from post facto special pleading? How much precision is possible in prediction, and does prediction have inherent limits? What obstacles prevent us from making predictions on the basis of intellectual resources already available, and what are the future prospects for prediction as a tool of applied sociology? Since a key to assessing a theory's predictive validity is its coherence with a wide-ranging and explanatorily integrated body of research, I will review how geopolitical theory has developed and how it meshes with the trend toward a state-centered and military-resource-oriented model of macropolitical change.

#### Development of the Geopolitical Theory of State Power

Geopolitical theory had its origins in the period when a distinctive line of conflict theory was first developed in Germany at the turn of the 20th century. Ratzenhofer and Gumplowicz stressed the military origins of the state; Ratzel's political geography discussed the propensity of large states to expand into continental empires. Weber's treatment of the development of the state was formulated in this context; for Weber ([1922] 1968, pp. 901–26), the dynamics of legitimacy as well as the formation of ethnic identification and nationalism are connected to military struggles among states and to the organizational mode by which varying proportions of the population are mobilized and equipped for fighting. For decades this aspect of Weber's theory lay dormant while emphasis was placed upon functionalist and cultural interpretations of Weber. Geopolitical thinking was in bad odor politically, associated with militaristic national policies advocated by early thinkers including Mackinder in England, Mahan in the United States, and Haushofer in Germany.

With the revival of conflict theory in the 1960s came the expansion of comparative historical research and renewed interest in the autonomous dynamics of the state. Geopolitics was rediscovered, separated from particularistic formulations, and pursued in a more analytical fashion (for overviews of the literature see Enggass [1986] and Hepple [1986]). In political science, classic geopolitical theory is generally regarded as foreshadowing the realist school of international relations. Let us consider the various contributions to modern sociological geopolitical theory as

they enter into the summary model that I constructed in 1978 (Collins 1978).<sup>2</sup> The principles are stated in terms of conditions for the expansion and contraction of the territorial power of states.<sup>3</sup>

PRINCIPLE 1.—*Size and resource advantage favors territorial expansion; other things being approximately equal, bigger, more populous, and resource-rich states expand militarily at the expense of smaller and poorer states.*

This principle is often stated in the literature on victory and defeat in war (Liddell-Hart 1970; Andreski 1971; Gilpin 1981; Modelski and Thompson 1988; Thompson 1988). Singer (1979) and Singer and Diehl (1990) find that the advantage is relatively small; but it is cumulative over time, insofar as marginally resource-dominant states add resources from their victims, while the latter grow cumulatively weaker.<sup>4</sup> This expansion may occur by direct acquisition and administration of territory. Resource-dominant states also expand by peaceful or quasi-peaceful means: by demanding that smaller client states contribute supplies or troops to joint alliances under central leadership and by adjudicating external and sometimes internal relations of weaker states. Through these mechanisms, the extent of de facto and often de jure military control over territory tends to grow.

PRINCIPLE 2.—*Geopositional or "marchland" advantage favors territorial expansion; states with enemies on fewer fronts expand at the expense of states with enemies on more borders.*

Here geography enters in two ways: natural barriers in the form of mountains, wide seas, and uninhabitable territories give some states ("marchlands") a "back wall" that enables them to concentrate their forces in fewer directions. On the other hand, large territories without natural barriers can support a multitude of states, especially if these are fertile agricultural land capable of supporting large populations. World

<sup>2</sup> Collins (1978) is the source for materials below not otherwise cited. The general model formulated in that article was based on synthesis of previous geopolitical theory literature, together with research using historical atlases covering 3,000 years of the Middle East and Europe, plus an independent series covering China. This basis is broader than that used by most of the international relations and geopolitical literature, which draws its evidence primarily from post-1500 Europe and secondarily from the classic phase of Greco-Roman antiquity.

<sup>3</sup> In contrast, one limit on much international relations literature is that it is concerned with explaining the incidence of war and the conditions for stability or transition in the interstate system. The broader issue, however, is to give the conditions determining the changing power of states and especially the power that results from wars.

<sup>4</sup> Boulding (1962, pp. 237–39) formally derived cumulative advantage in the case of two-person games.

historians (most comprehensively, McNeill [1963]) frequently note the pattern of marchland conquerors from the periphery of the great population areas. A high proportion of large-scale conquest states began in the marchland position; all seven of the unifiers of China following periods of multistate fragmentation came from those marchlands in the northern regions where population resources were relatively large vis-à-vis other marchlands. These cases point up the interplay between marchland and resource advantages. If there is an array of potential contenders situated in marchland positions, the one that will expand the furthest is the one that starts with the best local resource advantage and parlays that combination of advantages into a cumulative growth in resources as it expands against centrally located enemies.

PRINCIPLE 3.—*States in the middle of a geographical region tend over time to fragment into smaller units.*

This principle is an extension of the previous one. One reason that marchland states expand is that interior states are blocked, over the long run, from cumulative growth in their resource bases. Interior states have potential enemies and allies on many fronts; these situations foster balance-of-power diplomacy in which a defensive coalition forms against any momentarily dominant state (Morgenthau 1948; Gilpin 1981). Conflicts among interior states tend to become stalemated and thereby to eat up military resources without productive gain. Since such interior states are frequently located in good arable land, they often rank high in military resources; but they are structurally blocked, since their expansion possibilities are randomized, whereas only marchland states can enter into long-term cumulative expansion. Older comparative historians commonly misinterpreted the causes of this empirical pattern, attributing it to the greater vigor of barbarians from the periphery, against the decadence of advanced civilizations. When geositional advantages do not enter the issue, however, civilized high-resource states invariably win over resource-poor "barbarian" or tribal areas. The advantage of the periphery is not cultural but structural.

Further, there is visible in the historical atlases a pattern that goes beyond the mere advantage of marchland over center. Over time, in the absence of marchland conquest, internal areas tend to fragment into an increasing multiplicity of states: this pattern is found in China during several interdynastic periods, in Kievan Russia, in the Balkans after the decline of the Ottoman and the Austrian empires, and in the fragmentation of the medieval Holy Roman Empire into the *Kleinstaaterei* of Germany and Italy. Fragmentation occurs because interior states become militarily weakened, incapable of controlling secessions. Unstable and overlapping patterns of conquest and alliance divide administrative au-

thority and make the culture of political identification increasingly localistic.

**PRINCIPLE 4.**—*Cumulative processes bring periodic long-term simplification, with massive arms races and showdown wars between a few contenders.*

Principles 1–3 are cumulative. Big states swallow up the smaller or force them into alliances, and marchland states expand into the fragmented middle. The consequence is that over long periods of time (on the order of several centuries) the geopolitical situation undergoes drastic simplification.

This simplification may happen in a number of different ways. Historically one pattern is the growth of a single marchland conqueror that rapidly passes through a phase of accelerating conquest of interior states. This pattern is characteristic of geographically simple regions with only one major population zone, such as China or Mesopotamia in the period when it was an isolated region of agriculture. A second pattern, more characteristic of geographically differentiated western Eurasia after the diffusion of agricultural populations, is the growth of two rival marchland states, expanding into a stalemated central region from opposite directions. A third variant is simplification into two huge power blocks, one of them relatively more peripheral than the other.<sup>5</sup>

All of these situations create a period of high geopolitical tension: at minimum an intense arms race and diplomatic polarization, often culminating in a showdown war (in the terminology of international relations, a hegemonic war; see Gilpin 1981, pp. 186–200). In the first subtype mentioned above, the single large state continues its acceleration to effective conquest of its region. When there is a showdown between two large states or blocs, more possibilities hinge on their conflict. One state may destroy the other, opening the way to regional world conquest (Rome vs. Carthage in the western Mediterranean, opening the way to the victor's easy expansion). A historically common alternative is a stalemate between the two contenders, leading to disintegration of both; this disintegration can happen because of massive material losses on both sides in

<sup>5</sup> The last is the type of simplification that emerged from the multistate fragmentation of medieval Europe. A series of three large empires or coalitions have fought for hegemony: the Spanish/Habsburg empire vs. a coalition centered on France, French military and diplomatic expansion culminating in the Napoleonic empire; the German-centered axis vs. the allied periphery in World War I and World War II. The recurrence of this pattern, with different actors in the roles, suggests structural conditions such as those set forth in the cumulative dynamics of principles 1–3, plus randomizing factors that determine who plays which roles. In each case the most expansionary coalition fell to defeat through military exhaustion.

war or through the resource strains of lengthy arms races. In this case a bystander state has the opportunity to expand rapidly into the resulting power vacuum.<sup>6</sup>

**COROLLARY 4a.**—*Showdown wars generate the highest level of ferociousness.*

The wars fought at the point where regional world conquest was in the balance have been the most ferocious in terms of massacres of entire armies including captives and extermination of civilian populations. The rise in the level of deliberate destruction has been noted in the case of the Ch'in (first unifier of China after the prolonged Warring States period), the Assyrians (first unifier of Mesopotamia), the Romans in their war with Carthage, and the Mongols (first attempted conqueror of the entire Eurasian landmass); there are analogues in the drastic escalation of civilian casualties in the world wars of the 20th century. Gilpin (1981, pp. 200–201) notes the higher intensity of violence during hegemonic wars of Europe since the 17th century. Conversely, periods of balance-of-power diplomacy in regions of geopolitical fragmentation have usually been characterized by codes of honor restricting combat and limiting its damages. The mechanism connecting geopolitical conditions with the intensity of violence is the high level of emotional and ideological polarization in situations where drastic structural consequences hinge upon the outcome of battle; in contrast, where the fragmented structure of power is little affected by particular military outcomes and where balance-of-power negotiations lead to frequent switching of alliances, emotional polarization is low.<sup>7</sup>

**PRINCIPLE 5.**—*Overextension brings resource strain and state disintegration.*

The further that military power is projected from the home base, the higher the costs. At a calculable point, most resources are used up in meeting these costs; the results are increasing strain on resources at home

<sup>6</sup> Wars between the Roman and Persian empires culminating in the early seventh century left both exhausted in a power vacuum into which an Arab coalition suddenly expanded under the new ideology of Islam. Other cases, from Chinese history, are cited in Collins (1978). Exhaustion of Germany, Britain, and France in World War II left the world under the dominance of the two marchland states, the United States and Russia. A principal discussion in international relations literature (Waltz 1979; Gilpin 1981) concerns whether multisided balance-of-power situations or two-sided polarization leads to greater stability or to the likelihood of major war. The empirical materials cited are inconclusive, in part because of the relatively short time spans examined. Over periods of 250 years or more that I have examined in historical atlases, the geopolitical situation has typically simplified eventually and culminated in showdown war.

<sup>7</sup> The second part of this empirical pattern fits the Simmel-Coser principle: crosscutting ties reduce the intensity of conflict (Coser 1956, pp. 78–80).

and vulnerability to military defeat, raising the probability for rapid unraveling of military power.

The principle of overextension in terms of logistical loads has been widely noted in the theoretical and empirical literature. Stinchcombe (1968, pp. 218–30) and Boulding (1962, pp. 227–76) stated formal models for the decrease in military resources deliverable at increasing distances. Collins (1978) found that decline in virtually every case of the centralized Chinese dynasties was initiated by logistical strain and related defeats on distant frontiers. Luttwak (1976) showed the successive reductions in military resources accompanying Roman commitments on distant frontiers. Kennedy (1987) documented the overextension strains in the decline of the major European empires from the 15th to the 20th centuries. These case studies also show that the negative effects of overextension operate much more rapidly than the processes of cumulative growth in resources that fueled expansion; empires that reach the overextension point tend to lose control over military organization and political authority within a few years, resulting in regime downfall or state fragmentation.

Much (but not all) of the evidence for the overextension dynamic was based on historical comparisons among agrarian states; the same is true of principles 1–4. To settle the objection that principles based on these historical periods would be outdated under conditions of modern military and transportation technology, I examined the geopolitical stability of modern seaborne empires and the military effectiveness of air power (Collins 1981). The logistical costs and vulnerabilities of these technologies offset their hypothetical increase in range;<sup>8</sup> accordingly, the basic geopolitical principles have remained valid.

The preceding is the theoretical base from which I made a prediction about the future of Russian state power. (For a schematic description of the geopolitical model, see fig. 1.) In the empirical and theoretical literature, the principles of resource advantage, marchland advantage, and overextension are well documented. There were abundant historical materials about showdown wars but no general theoretical formulation. The cumulative nature of some of these processes had been noted. The

<sup>8</sup> That is to say, modern technology makes possible a large increase in the distance at which military force can be projected for short periods of time. Against sustained resistance (even of a much lower military resource level), however, logistical costs at long distance are determinative, as the United States found in Vietnam during 1963–75. In the long run of the past 2,000 years, there has been only a modest increase in the maximal size of empires and no tendency to increase the frequency of approaches to world empires. This fact bolsters the inference that the ratio between transportation range and logistical cost has not changed greatly. Logistical constraints on the movement of effective military power under various historical conditions have been examined by Van Creveld (1977, 1991)

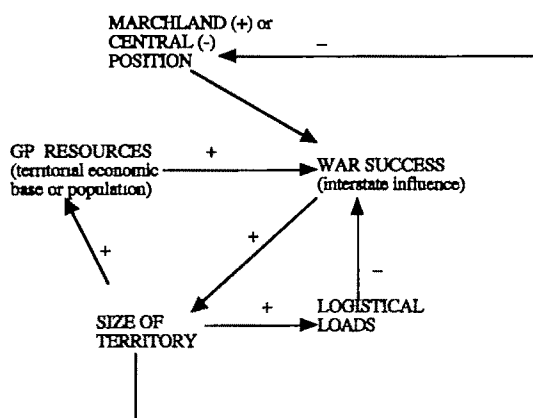


FIG. 1.—Geopolitical model

theoretical contribution of Collins (1978) was to sharpen the formulation of the fragmentation of the middle (as the converse of the marchland advantage) and to connect the entire package of principles as a dynamic of mutually reinforcing processes, leading periodically to long-term geopolitical simplifications and turning points.<sup>9</sup>

### The Connection of Geopolitical Theory with General Theory of State Formation and State Breakdown

Geopolitical theory was for long a minor interest in sociology; its importance has risen as the theory of the state has become more central. One trend in macrosociology since the 1960s has been emphasis upon the external relations among social units. World-system theory (Wallerstein 1974–89; Chase-Dunn 1989) has focused upon the dynamics among zones of the world economy. Since hegemony in the world system is based

<sup>9</sup> Formally stated, the five principles may be combined into a single, complex expression. Marchland advantage is weighted by the relative resource levels of adjacent states; overextension is the fundamental principle for stating the relative vulnerability of particular geographical points to states with given resources and logistical loads. Showdown wars result from long-term cumulation of resource and geopotential advantages and disadvantages. The technical difficulty in investigating this formalism is that mathematical and computer simulation methods are straightforward only in calculating the geopolitical strength of a single state with arbitrary values for its external environment. There is no simple way to model the general pattern of systems of multiple states in topological space, (i.e., the variety of possible spatial configurations that affect the marchland advantage). So far computer simulations have been carried out only on simplified versions of geopolitical principles; see Hanneman, Collins, and Mordt (1995).



on mutually reinforcing economic and military dominance, autonomous geopolitical conditions need to be inserted in the world-system model to account for shifts in hegemony and the rises and falls of states within the world system. In another externally oriented approach, Bendix (1967) formulated the process of modernization, not as parallel and indigenously driven developments, but as a chain of emulation among successive "leader" and "follower" states. Bendix's model includes some Weberian inspiration, since the basis of emulation is the power-prestige of states in the interstate arena, and this power-prestige may be construed as primarily geopolitical dominance. Bendix's external starting point for internal state changes was adopted and extended by Skocpol (1979) in her theory of revolutions initiated by military costs.

By the 1980s, there was a full-scale movement for "bringing the state back in" (Evans, Rueschemeyer, and Skocpol 1985). The agenda was to formulate a theory of the autonomous dynamics of the state, which can of course interact with economic, cultural, and other dynamics but is not reducible to them. The general issue may be conceived of as formulating the dynamics that determine state growth and state decline. Each of these directions of change contains subdimensions. States grow intensively (in their internal organizational size and capacity of control) and extensively (in territory); they go into crisis or break down in a number of ways (again, organizational breakdown and loss of extractive capability on one hand; as well as territorial contraction and disintegration on the other). The theory of revolutions is the topic that has received the most attention. We should keep in mind that revolution is one segment of the continuum of state breakdown, and breakdown in turn is the obverse of the more general question of state formation. Viewed in overarching perspective, what we are concerned with are the conditions that move states in various directions along these continua.

Tilly (1990) provides a recent summary of the upward pathways of state formation. The core of the state is its military organization, together with the administrative apparatus for extracting economic resources to support it. Once this apparatus was in existence, it could be used for other purposes as well (including economic regulation and infrastructure, welfare, and cultural dissemination); this part of state organization became relatively large in most instances only very recently, building upon the core of organization servicing the military. Throughout most of history, the state's expenses were primarily its current military forces plus debts incurred from previous warfare (Mann 1986, pp. 416–46). The "military revolution" of 1500–1800 (Parker 1988) escalated the size of military forces, along with their costs of equipment, their permanence, and the centralization of their control. So far we have a military dynamic driving the intensive growth of internal state organization over time.

Tilly (1990) shows that variations in the forms of state organization are explained by conditions affecting the kinds of resources that states can draw upon to support this military expansion. Depending upon whether concentrated sources of capital (mainly urban-trade economies) or dispersed economic resources (agricultural land) were relatively more available within their territories, states followed different trajectories of organizational and military growth. At one extreme, rulers rented short-term military force in collaboration with capitalists, thereby fostering shared power in urban oligarchies and federations and laying down structural bases for republics. At the other extreme, extensive landed conquest was the route to growth of states dominated by military aristocracies. The mixed form where both resource bases were combined, Tilly finds, was the pathway to the centralized nation-state, which eventually forced the alternative state structures to emulate it because of its superior capacity for military mobilization and intensity of administrative control. An important variation among state structures is brought out by Downing (1993), who analyzes the effects of the military revolution upon democracy. The collegial structures of power-sharing in medieval states were eroded by the rise of centrally provisioned armies; the states most vulnerable to this erosion became autocracies, whereas those that most delayed the administrative structures of the military revolution developed the strongest parliamentary-representative institutions.

### The Mechanisms of State Breakdown

Schematically, the structure of military resources has been the key to the internal forms of state organization. Connecting such resources to geopolitical theory are feedback loops between the two spheres. Geopolitical power is in part the result of internal resource extraction (via principle 1); in turn the rise or fall in overall geopolitical position (as the result of all five principles) feeds back into the amount of territorial resources available and the rate at which those resources are eaten up in war preparations and violent destruction (see fig. 2).

Consider now the negative side of the state-formation continuum. In recent decades of research, state-centered models have become prevalent in the theory of state breakdown and revolution. A key is the vulnerability of states and their rulers to crises of resource extraction, relative to state costs. Skocpol's (1979) pioneering formulation might be called the material economy of the state; the state itself is an economic entity, which forms a class interest in its own right. Primary among such interests is the state administrative class, whose economic (as well as power and prestige) interests favor the expansion of extractive capability. The principal class of opponents is the propertied elite (in agrarian societies, the

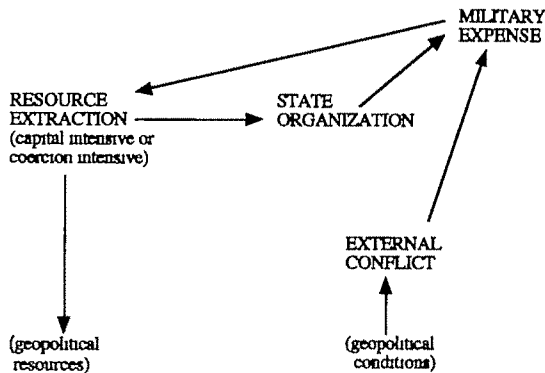


FIG. 2.—State growth model. Tilly

landowning nobles), whose interests are in evading extraction from their own resources. Insofar as these two classes are intertwined socially and politically, under conditions of state budgetary crisis, conflicts break out within the elite. These conflicts, together with the financial aspect of the crisis itself, which paralyzes or alienates the military forces, culminate in full-scale breakdown of the state at the top. This breakdown opens the way to revolutionary forces from below.<sup>10</sup>

The outstanding test and theoretical elaboration of the model of state breakdown is Goldstone (1991); this work builds lengthy time series of empirical indicators of the several aspects of state strain, showing that the composite index of pressures for crisis corresponds with historical ups and downs of state crisis. In both Skocpol and Goldstone, state breakdown results from the combination of (a) state fiscal strain, (b) intrastate conflict that paralyzes the government, and (c) popular revolt. Skocpol stresses military strain as the prime source of state fiscal and administrative crisis; the sources of such military strains, further along the causal chain, are specified by geopolitical theory. Goldstone adds causal paths to

<sup>10</sup> I leave aside here consideration of the various kinds of revolutionary coalitions and countercoalitions, since these are to a large extent outside the realm of the state. In this part of her model, Skocpol draws upon Moore (1966). Moore's analysis of the varying structural consequences of different types of class relations in capitalist agriculture have been tested on a further range of historical cases by Paige (1975); it converges as well with the early formulation of Stinchcombe (1961). Moore revives Marxian class analysis of revolutions by shifting the focus from industrial to agricultural property relations; Skocpol opens another direction by adding a focus upon the autonomous dynamics of the state. These lines of theory are not necessarily alternatives. The state breakdown model is crucial for the initiation of revolution; the actors in the conflict and the directions that it takes, especially after the revolution, are affected by economic property relations. The relative weight of these vis-à-vis other factors remains to be clarified by future research.

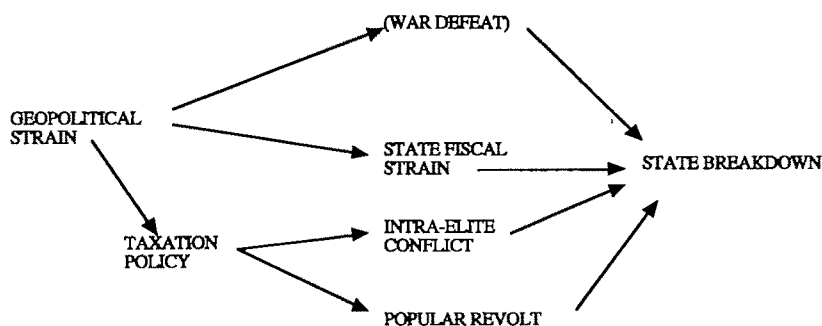


FIG. 3.—State breakdown model: Skocpol

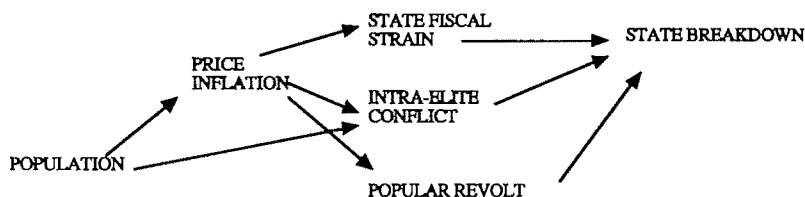


FIG. 4.—State breakdown model: Goldstone

all three aspects of state crisis, focusing on the ways in which population pressures, mediated by prices, inflation, and taxation, affect conditions. *a-c*. Goldstone argues against the Skocpolian emphasis upon military sources of state fiscal strain, but the two causal chains are not mutually exclusive; in his own model, the key to state breakdown is not population pressure per se but the relative overall balance between state obligations and state resources (see figs. 3 and 4).<sup>11</sup> In cases where military expenditure and past military debt are the bulk of the state budget, geopolitical strain should generate strong pressures for state breakdown, whether or not this is the exclusive source of such pressures. In more extreme cases, state breakdown has been precipitated directly from disintegration of the military apparatus in war.

Disputes over the emphasis given to particular parts of the causal chain should not obscure the cumulative achievement of this series of studies. We have good evidence for "a core model of state breakdown—fiscal/administrative strain, elite conflict, popular revolt—plus a number of

<sup>11</sup> A key comparison is the Meiji Restoration (Goldstone 1991, pp. 404–14), during which a period of population stability rather than growth played into state fiscal crisis and other aspects of strain; this situation occurred because incomes were fixed in kind while cash prices rose with economic expansion. For a more detailed comparison of the Skocpol and Goldstone models, see Collins (1993).

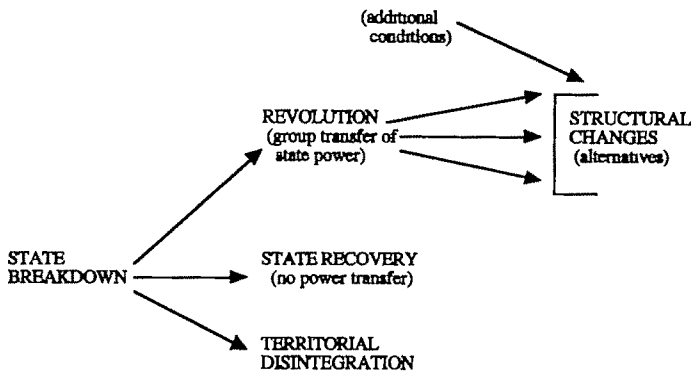


FIG. 5.—State breakdown and revolution

pathways towards crisis conditions in these factors. Population growth can sometimes play a very large role in building up crisis; at other times GP conditions have overwhelming effects. In many cases, population and geopolitics interact" (Collins 1993, p. 121). For theoretical utility, we would not want to narrow the application of the core model to historical periods when population growth is a major driving force in the background variables. The advance of theory is just such a development of a core model, with ancillary models that make it applicable to a variety of historical conditions.

Our understanding of the relation between state breakdown and revolution theory is now becoming clarified (see fig. 5). As Goldstone shows, not all state breakdowns are followed by revolutions in the specific sense of wholesale transformation of the ruling elite accompanied by political and economic restructuring. More specific theories of revolution (and other postbreakdown paths) are needed if we are to understand and predict these kinds of consequences. It should be noted that my analysis of Soviet decline predicted breakdown; it did not offer a theoretical basis for predicting what kind of regime would follow. As of 1994, it remains undecided whether the former Soviet Union will indeed undergo revolution in the full sense of the term.

#### Legitimacy as a Variable Affected by Geopolitical Power-Prestige

My own version of the tie-in between geopolitical principles and state breakdown stresses the mechanism of legitimacy. This is not to slight the Skocpol-Goldstone model of administrative crisis through fiscal strain, intraelite conflict, and popular dissidence, but to add to them a processual dynamic that ties directly into the emotional level of political maneuver-

ing. It should be stressed that legitimacy is not something to be viewed as an abstract and constant property of the political system. On the micro- and mesolevel of social interaction, legitimacy is directly connected with solidarity and loyalty among political groups and with the enthusiasm or acquiescence of the masses; conversely, delegitimation is an emotional and cognitive condition that prevails when elite political activists are divided and uncertain and when masses move from alienated disaffection to acts of opposition. Weber's legacy has misled us here; so much attention has been placed upon the typology of traditional, rational-legal, and charismatic legitimacy as a static classification device that the processual aspects of legitimation are slighted.

Legitimacy is a variable on two levels. There is a continuum of popularity of particular political leaders. There is considerable evidence, from opinion polls in modern times and from historical accounts for earlier periods, that political leaders' popularity is most strongly affected by periods of military conflict (Ostrum and Simon 1985; Norpoth 1987; Bueno de Mesquite, Siverson, and Woller 1992; Gallup polls cited in Hanneman, Collins, and Mordt [1995]). On this level, we have evidence of the link between the military power-prestige of the state and the legitimacy of its rulers. Militarily successful rulers, favored by geopolitical circumstances, make their own legitimacy when they impose their power at home, even if they may have started as coercive and illegitimate usurpers. And war-defeated leaders, however legitimate they may have been, run increased risks of being ousted. Furthermore, geopolitical conditions determine not only the degree of personal legitimacy of rulers along this part of the continuum; extreme geopolitical crises are the prime cause of delegitimation of the entire institutional order. Empirically, the continuum extends from degrees of personal popularity, through legitimation or delegitimation of the political structure. This is so at either end of the continuum: extremely popular individuals (who would be called charismatic) are able to extend the aura of emotional mobilization to legitimating an entirely new kind of state; at the low end, beyond intensely negative levels of rulers' personal popularity comes the delegitimation of the entire state.<sup>12</sup>

<sup>12</sup> Although this claim has not been tested systematically, historical evidence would seem to support it. It is difficult to cite any instance in which an entire political (or religious) order became delegitimated and overturned, while the ruler remained personally popular; conversely, there appear to be no instances in which a new institutional order was instituted by a leader who was very unpopular. The distinction between personal and institutional legitimacy appears to operate empirically only in the moderate levels of the continuum. This is not to say that intense personal unpopularity of the ruler is the only pathway to institutional delegitimation (an aspect of what we commonly call revolution). In some instances this is true: personal dislike of

The argument does not hinge upon the popularity of individual rulers per se (although such popularity is a convenient source of evidence for the connection between legitimacy and geopolitics). Individuals are caught up in larger processes. The swings of interstate power-prestige, above all other factors, bring about rapid shifts in the legitimacy of rulers. Moreover, a geopolitically induced condition of state administrative crisis, together with inraelite conflict, tends to bring about the rapid alternation of leaders in office; if this reaches a sufficiently high level of infighting and chaotic turnover, not only do particular leaders lack respect, but the very institution of leadership appears ineffectual. It is through these channels that macroprocesses of geopolitics become translated into the specific events and personalities of situations of state crisis and breakdown.

#### Geopolitical Theory's Prediction of the Russian Empire's Breakdown

Consider now the specific predictors of the Soviet collapse and how they worked out.

*Size and resource advantage* had been building up since the expansion of the original small state of Muscovy in the 1390s; by the late 1700s Russia could field the largest armies in Europe.<sup>13</sup> By the mid-20th century this advantage was reversed; Russia and its allies became outweighed by their enemies in total population by a ratio of 3.5:1 and in total economic resources by 4.6:1. In troop numbers, the two blocks were much closer: Russian enemies led by 1.7:1 in active troops and by only 1.1:1 when reserves are included. Thus the Soviet bloc was 3.5 times as heavily mobilized, in the ratio of troops to population base, as its enemies. Correspondingly, the Soviet Union spent a disproportionate amount on its military budget (as much as 20% of the GNP), at the expense of civilian expenditure. It was this strain that Gorbachev addressed early in his period of reform, from 1985 onward announcing and to some extent implementing plans to scale back military forces and convert to civilian production (Becker 1986, 1987; Bernstein 1989; Gelman 1989).

The *marchland advantage* and *fragmentation of the interior* had

---

the religion and the sexual morality of King Charles I was a prelude to the English revolution of 1640; swings in the personal popularity of Gorbachev were part of the dynamic of delegitimation of the Soviet regime. On the other hand, victims of revolution like Louis XVI in France and Nicholas II in Russia were not especially hated, although they were not popular figures (Stone 1979, p. 335). In all these cases, the halo of regime crisis accelerated personal unpopularity.

<sup>13</sup> Data here and in the following are from Collins (1986), *Facts on File* (1984–91), and *Keesings' Record of World Events* (1984–91) unless otherwise specified.

worked to Russia's advantage during its period of expansion. Moscow originally expanded from the "back wall" of the sparsely populated northern forest zone into the fragmented states of the central Russian planes and the disintegration of the Mongol empire. The fragmentation of the interior state of Poland, attacked from three sides, brought Russia to a stable frontier in Europe, until the defeat of the German central power in World War II in another multifront war allowed further expansion of Russian territorial control in the form of its satellite empire of Eastern Europe. Russian expansion into Siberia in the 1600s had taken place against sparse tribal populations; in south and central Asia and the Caucasus, expansion went on into the late 1800s against petty states left by the disintegration of a succession of empires based in Anatolia or Iran. These geopolitical conditions were reversed to an increasing degree after 1900 (i.e., the negative feedback loops in fig. 1 were now becoming predominant over the positive loops). Expansion to the Far East brought Russia up against Japan and defeat in the war of 1904-5 (thereby setting off an abortive revolution at home); with the reconsolidation of a Chinese state after 1949, fighting again took place against China in 1969 (following battles in 1929 and 1945), and heavy troop placements were maintained on these borders. In the south and west, the former buffer zones of small states were eaten up, and Russian troops were maintained in forward positions directly confronting the forces of the NATO alliance. By the 1950s, the Russian empire defended a land frontier of 58,000 kilometers.

My prediction was that the coincidence of crises on several frontiers at once would precipitate the unraveling of the empire. The logic is similar to that which Perrow (1984) refers to as "normal accidents" in complex organizational systems; in such structures, local breakdowns that can be remedied in isolation ramify into a systemwide crisis when they occur simultaneously. Such events occur on a probabilistic basis, increasing with the number of components. Russia had faced a series of uprisings in its Eastern European satellites (in 1953, 1956, and 1968), which it put down with the aid of loyal Warsaw Pact troops. But the ability to command loyalty depends upon perceptions of overall coercive capability, and thus is subject to a "tipping phenomenon" (Schelling 1962, pp. 51-118) when an atmosphere of crisis makes sanctions for noncompliance seem unlikely. My projection was that the interaction of geopolitical disadvantage with overextension and with enemy local resource advantage would bring a combination of failures in far-flung regions (such as military interventions to prop up client states like Afghanistan) and would eventually reach a tipping point.<sup>14</sup>

<sup>14</sup> "Thus it is highly likely that, once a first round of serious crises caused the loss of Eastern Europe or other distant territory, there would be set in motion cumulative processes of internal weakening, culminating in the eventual loss of the next tier of



What happened in fact was stalemate and growing domestic disgruntlement over the Afghanistan intervention begun in 1979, leading up to the withdrawal of Soviet troops beginning in 1988. The military-expansionist regime that prevailed up to the time of Brezhnev's death in 1982 and his successors Andropov and Chernenko was replaced in 1985 by a reform faction under Gorbachev (a former Andropov protégé). Nationalist agitation within the Soviet Union began immediately, from 1986 onward. In the outer periphery of Caucasian and central Asian republics (including Uzbekistan, adjacent to the Afghanistan guerilla war) interethnic violence broke out in 1988 and 1989. Meanwhile, mobilization of dissidence in the west suddenly went unchecked. Dissidence based on trade union organization had developed in Poland in 1980 and had been repressed by martial law late in that year, following the previous pattern of rebellion and defeat in the East European satellites. In the context of the unfolding geopolitical crisis, however, in 1988 officials of the Baltic republics of the Soviet Union backed proposals for complete autonomy. In Poland the Solidarity union strikes led to reform concessions in 1988 and to free elections in June 1989; the victorious Solidarity candidates were invited by the Communist regime to form a coalition government in August. Simultaneously in Hungary, a split within the ruling Communist Party transformed the unified autocracy into shared rule in June 1989.

Ethnic population movements across borders were initiated under Gorbachev in the policy of allowing Jewish emigration to Israel; beginning with small numbers in 1986, this emigration accelerated to 20,000 in 1988 and 60,000 in 1989. The era of openness let loose pressures for ethnic migrations all across the Soviet bloc. Within the Russian Federation itself, people dislocated during the Stalin years now sought to return to ethnic homelands. In 1987 Hungary began to grant asylum to ethnic Hungarians fleeing Romania, resulting in Romania temporarily sealing the border then opening it under pressure in 1989. In the same year Turkey sealed its border with Bulgaria to stem the flow of some 300,000 Bulgarians into Turkey. These developments encouraged a similar movement for exit visas by East Germans through Hungary to Western Europe in the late summer and fall of 1989. It was this movement that proved to be the tipping point; when Czechoslovakian borders, under pressure of mass movement agitation, were reopened on November 1, 1989, the flood of refugees turned into open regime opposition that mounted an attack on the symbolic border, the Berlin Wall.

In the two weeks of November 9–24, 1989, confrontations with regime forces in East Germany and Czechoslovakia suddenly resulted in conces-

---

ethnically distinct conquest. the Baltic states, the Ukraine, the Caucasus, and the central-Asian Moslem territories" (Collins 1986, p. 203).

sions, as Gorbachev refused to commit Soviet Warsaw Pact forces to controlling dissidents, in effect overruling the use of repressive force. The mutual threat that held together discipline within the Eastern European armed forces now disintegrated; its disappearance was rapidly followed in two months by uprisings and transfers of power all across the Eastern European block. These events at the outermost periphery of satellites accelerated the disintegration of the next tier of the Russian empire. In March 1991 Gorbachev's plans for liberalization of the Soviet Union into a looser federation were upstaged by the withdrawal of six of the republics on the westernmost peripheries (the Baltic republics plus Moldavia, Armenia, and Georgia). Following the breakdown of authority in August 1991 with the failed coup attempt of the anti-Gorbachev faction, the entire Soviet Union formally disintegrated. The pattern, as predicted, was a coincidence of crises on multiple fronts, interacting and accelerating past a tipping point into a generalized breakdown of territorial authority.

*Showdown through massive arms races and wars and an atmosphere of military ferociousness* were in evidence since midcentury. The world geopolitical situation had been simplifying drastically during the 20th century. The Nazi empire and its opposing coalition marked one approach to all-out confrontation of opposing blocks; the Soviet/NATO confrontation absorbed and divided the resources of the defeated party. In this light, the nuclear arms race, with its potential for the most massive destruction of civilian population in world history, does not appear such an anomaly. It fits the pattern of the highest level of military ferociousness, which is associated with showdown wars when cumulation of geopolitical resources reaches the point of maximal simplification and polarization. This part of geopolitical theory also supported the prediction of Russian power loss, although it left two subpaths: (1) mutual resource exhaustion through arms race or (2) open war.<sup>15</sup>

The historical outcome was that Gorbachev's reform movement, soon after taking office, began to unilaterally reduce nuclear armaments and negotiated bilateral reductions following Gorbachev's 1987 summit meeting with President Reagan. From the point of view of geopolitical prediction, resource exhaustion by showdown confrontation had set in. The nuclear arms race had accelerated in costs drastically during the early 1980s. The proximal cause was the arms buildup in the United States under the Reagan administration. From the point of view of conflict

<sup>15</sup> A third subpath, victory by one state leading to world domination, was ruled out because nuclear war would cause extreme resource destruction (Office of Technology Assessment 1979). In addition, principles 1–3 and 5 indicated that Russia would not be the victor in a war with the entire enemy coalition.

theory, this escalation in the emotional tone of confrontation followed from a series of reciprocal threats between the United States and Soviet Union since the late 1940s. At some point the mutual escalation must result either in nuclear war or in exhaustion and deescalation. Nuclear armaments costs combined with Russia's other military expenses to cause the resource crisis of the 1980s that brought the Gorbachev faction to power. Gorbachev acquired his initial public charisma by reversing the nuclear escalation; in the period 1986–88 he had enormous prestige, especially in Western Europe, on the basis of this policy. This popularity of Gorbachev in turn first broadcast his image as a liberalizing reformer, and his travels abroad on behalf of deescalation encouraged Eastern European dissidents.

Here too we find an interaction among the several processes. The threat of immediate use of nuclear weapons by NATO was a policy adopted to countervail the threat of Warsaw Pact forces close to the population centers of Western Europe; in return, the Soviet rationale for maintaining those forces in Eastern Europe was to deter a nuclear strike. At the same time, the presence of Warsaw Pact forces served to bolster satellite regimes. Hence the commencement of nuclear disarmament carried the expectation of troop withdrawals, and this expectation in turn undermined the mutually monitoring coercive coalitions that constituted the authority of the satellite regimes.

*Overextension* had plagued Russia's geopolitical position in east Asia since the beginning of the 20th century. In the west, the acquisition of the satellite empire after 1945 added overextension in another direction, as did intervention in Afghanistan on the central Asian frontier. There are two aspects of overextension: (a) logistical costs that eat up increasing proportions of military resources in transportation and (b) ideological or cultural resistance that manifests itself in ethnic hostility. The logistical overextension involved a combination of distant military commitments: the Soviet Union's frontier itself; the additional effort to build world naval power, which was at its height in the early 1980s with the creation of a massive carrier fleet for Pacific, Arctic and North Atlantic, and Black Sea and Mediterranean ports; plus military aid transported to allies in Cuba, Vietnam, the Middle East, and Africa.

The second aspect of overextension results because far-flung conquest results in multiethnic empires. In a wide range of historical instances (Collins 1978) ethnic hostility to foreign rule was mobilized increasingly when an empire controlled two or more layers of ethnically distinct territories beyond the home ethnic heartland (its inner zone of relative ethnic homogeneity). In this respect, Russian overextension approached highly dangerous levels in the period after 1945. In the satellite zone of Eastern Europe, Soviet troops upheld power in territories two and three ethnic

layers distant from the Russian homeland. In Afghanistan and in the Caucasus, where Russian troops were stationed against Turkish and Iranian forces, force was again projected beyond a tier of non-Russian ethnic territories. The dissidence mobilized in these regions, both in the period 1953–68 and with greater effect in the breakup of 1988–91, drew primarily upon ethnic identities and hostility to locally stationed Russian personnel.

I will discuss later the extent to which ethnic identifications were responsible for the Soviet breakdown. Here I want to stress the point that ethnic hostility is differentially mobilized and occurs first and most intensely in regions that are geopolitically vulnerable to breakaway. The outer tier of the Russian empire, the satellite states, were less strongly integrated organizationally; national government structures were left in place along with local-language dominance. Movements mobilized to take advantage of Russian geopolitical weakness undercut local Communist authorities oriented toward the Russian alliance; consequently, satellite breakaways tended toward regime revolutions. Within the Soviet Union itself, the movements that led to formal breakup were more easily mobilized because the component republics were already organized as nominally sovereign ethnic groups (Waller 1992). This organization resulted from the Russian empire and its successor, the Soviet Union, being multiethnic conquest states. Multiple ethnicities were embodied in the structure during both its growing and its declining phases; thus ethnicity cannot be considered a primary cause of the Soviet breakdown but is, rather, the organizational medium through which geopolitical overextension worked itself out.

*State breakdown and the accelerating legitimacy crisis* unfolded along the lines of geopolitical theory. The cumulation of geopolitical strains brought the Soviet Union in the mid-1980s to the point of *state fiscal crisis*. Political breakdown was first manifested in *intraelite conflict*. The power-prestige of the military-expansion faction was reduced by its failure to keep up with U.S. armaments and by its debacle in Afghanistan. The Gorbachev reform faction came to power and became embroiled in conflicts with the faction based in the Soviet military-industrial complex. This split is structurally analogous to the elite factionation depicted in the Skocpol-Goldstone model: a faction concerned with state fiscal health (pure state-class interest) versus a faction of "aristocrats" whose material property is upheld by the state and who shunt the burden of extraction onto other classes. Gorbachev threatened to cut back subsidies to the military-industrial sector, hitherto the strongest part of the Soviet economy. The result was the equivalent of an aristocratic "tax revolt," as the military-industrial sector sabotaged economic reforms in the direction of conversion to a civilian economy. As a further result, the reform period

became one of worsening economic crisis. Gorbachev was in the unenviable transitional position along the slippery slope of reform under incipient breakdown conditions, analogous to the Necker ministry in France in the 1780s. The worsening geopolitical, economic, and political situations reinforced one another. Gorbachev's reform faction was unable to muster resources to bolster its own power. Ultimately, the continuing intralite conflict, coming to a head in the attempted coup by the military-industrial faction in 1991, broke apart state authority and opened the way to revolution in a fashion paralleling the aristocratic revolt that precipitated the downfall of the French monarchy in 1789.<sup>16</sup>

The third component of the state breakdown model is the mobilization of oppositional class forces from below (*popular revolt*). The greatest variability in the general theory is at this level. In place of peasant revolts, which fit this slot when the model is applied to agrarian-capitalist societies, we find a combination of dissidents centered initially on the nominally ethnic organizations in the peripheral republics, joined at the height of disturbance by the educated populace in the main urban centers (Bessinger 1990; Sedaitis and Butterfield 1991; Roeder 1991). Here again there was a cumulation and acceleration of the conditions of breakdown and dissident movement mobilization.

Gorbachev's faction, in its efforts to create support in the intralite conflict against the military-industrial faction, carried out a series of political liberalizations. Their early maneuverings attempted to undercut the power of officials from the Brezhnev era by attacking state bureaucracy and the Communist Party itself through calls for self-management and accelerated economic development. The idea of *perestroika*, introduced in April 1986, can be interpreted as an effort to broaden the base of Gorbachev's support for redirecting state activities from military mobilization to economic development. Political dissidents were released from prison or freed from internal exile in 1986 and 1987; popular assemblies were developed to undercut existing Communist Party organization. A series of reforms in the electoral process, beginning in 1988, created openings in the union republics, culminating in the first openly elected USSR Congress of People's Deputies in 1989. The growth of these alternative structures—what Trotsky in his theory of revolution called “dual sovereignty”—provided an organizational base upon which the legitimacy of

<sup>16</sup> This is not to say there were no other sources of intralite conflict going back before the period of geopolitical crisis in the 1980s. As noted below (in the section “Capitalism versus Socialism”), issues of economic inefficiency and market-oriented reforms also split the Communist elite; but comparisons among China, the Soviet Union, and Eastern Europe indicate these economic issues cannot have been the key to the Soviet empire's breakdown.

the entire institutional order could be called into question. Initially, the mass mobilization encouraged by Gorbachev served to make him into a charismatic figure—the speechmaker amidst the admiring crowds—shifting legitimacy onto him personally. But personal legitimacy is a two-edged sword, especially in a situation of volatile and shifting resource bases. With the deepening of structural crisis and Gorbachev's failure in intraelite conflict, he was suddenly delegitimated; in turn, his personal fall delegitimated the entire regime.

## THE ISSUE OF PREDICTION IN SOCIOLOGY

### What Counts as a Valid Prediction?

Can successful historical predictions be made? Obviously they can. But there is a difference between a sociological prediction and a guess or wishful thinking. A valid prediction requires two things. First, there must be a theory that gives the conditions under which various things happen or do not happen (i.e., a model that culminates in if-then statements). This standard of theory is more stringent than what sociologists generally mean by the term. It is not a category scheme, nor a meta-theory, nor even a process model that lacks observable if-then consequences. Second, there must also be empirical information about the starting points, the conditions at the beginning of the if-then statement. My prediction of the Soviet collapse was based on both the principles of geopolitical theory and empirical data about the condition of the Soviet Union and its enemies as of the 1970s.

A good deal of confusion in the debate over whether prediction is possible has come from failure to distinguish the two components. In the absence of a theory, prediction is merely empirical extrapolation. This does not give very much vision if the extrapolation is short-term, when all that one is doing is extending a process already under way without knowing what causes it to change direction. Long-run empirical extrapolation can be notoriously unreliable. Much of the prediction that sociologists do is of this sort. For instance, we hear statements that by the middle of the 21st century more than half of American children will be minorities. In the absence of a theory of what determines ethnic identities, this sort of prediction is questionable, for it assumes that there will be no ethnic assimilation nor change in the social categories of ethnicity. The most famous failure in sociological prediction was made in 1940 by demographers who, lacking any theory that could have encompassed a baby boom, extrapolated trends and made population projections that turned out 100 million short.

Theoretical principles plus empirical data are necessary for prediction in which we can have some confidence, prediction that is more than a

guess. This implies that the theory must also be validated. The validity of a macrosociological theory is never an all-or-nothing matter; in dealing with complex processual models, with feedbacks among both internal processes and external relations among such units as states, there is never a simple and clear-cut statistical acceptance or rejection of an overall model. That is not to say that particular components of the model cannot be falsified by data; but resisting falsification in piecemeal analysis is not the primary way in which the plausibility of a macrodynamic model is built up. The success of my Russian empire prediction adds to our assessment of its validity, but if the theory had no other basis than this case our confidence in its further applicability would remain rather tentative. Overall coherence among all the sources of evidence is central for our judgments of validity, and that coherence is displayed to the degree that the theoretical statements summarizing various cases can be made logically consistent with one another. The coherence among geopolitical theory, the military resource theory of state formation, and state breakdown theory is a source of mutual validation for all. In the foregoing review of these models, I have cited a body of research, some of which was done since the time I put forward in 1980 the geopolitical prediction of Soviet collapse. Research and theory development along these lines continues at present. For validating the overall model, and thereby showing that the basis for prediction was systematic rather than ad hoc, the time at which such research is done is irrelevant. In this sense, the ongoing cumulation of macropolitical theory continues to build up (or conceivably may undermine) the validity of the geopolitical prediction.

From this point of view, let us consider some of the other predictions and after-the-fact explanations of the collapse of the Soviet empire.<sup>17</sup>

*Ethnic revolt.*—The most notable prediction with which I was familiar in 1980 was made by d'Encausse (1979). Her prediction was based on an empirical extrapolation. D'Encausse calculated population trends for Russians and other ethnicities of the Soviet Union and concluded that the empire would disintegrate in the 21st century as non-Russian nationalities became the majority. Was this a valid basis for prediction? Can we extract a theoretical principle from d'Encausse's use of ethnic population trends? D'Encausse's prediction implies that sheer size of ethnic groups determines relative political strength and implicitly that multiethnic states tend toward rebellion and disintegration. What is missing is a broader theory of the conditions under which ethnic groups assimilate, remain distinct, or for that matter fragment into still smaller subethnici-

<sup>17</sup> See Collins and Waller (1992) and Gaddis (1992) for more extensive review of theories that mispredicted the Soviet collapse.

ties. But trends in ethnic boundaries are highly variable, and the major determinant seems to be geopolitical (Collins and Waller 1994). That is, when the core state is geopolitically strong, the prestige of its dominant ethnic group is high and it becomes the target for ethnic assimilation; it is when geopolitical disintegration is already under way that ethnic separatist movements are mobilized as vehicles for the decentralization of power. The conclusion follows that, in the absence of the package of geopolitical conditions, predictions of relative ethnic population size would not have been a valid prediction of Soviet collapse.<sup>18</sup>

It is important not to raise these issues merely in the spirit of determining who should get credit for making the right prediction. Prediction is not a one-shot deal, picking out the lucky hits from a series of misses; it is useful only if we know that we have a reliable instrument, a tool we can use repeatedly and in various circumstances. It is a crucial question, for instance, for the situation of Eastern Europe and Central Asia in the 1990s, and doubtless of the 21st century, whether ethnic population size alone drives state dissolution or whether overall geopolitical conditions determine the direction of ethnic assimilation or disassimilation. Only on the basis of a good general theory can we offer predictive insight.<sup>19</sup>

In this spirit, let us consider the post facto claims that have been made after the 1989–90 revolutions and especially after the summer of 1991. Most of these theories fall into two categories: (a) oppressive states like the Soviet regimes were bound to be thrown off by their people, and (b) centralized (and a fortiori socialist) organization of the economy inevitably loses out in competition with the higher productivity of free market capitalism. Leave aside the fact that virtually everyone before the late 1980s, including most social scientists, regarded Soviet socialist regimes as essentially permanent. How do the two popular retrospective claims stand up as theories?

*Overthrow of oppression.*—Theory *a* is clearly wrong. It is a battle cry of exultation at the moment when the statues were toppling. But the oppressiveness of a state has been a poor predictor of its vulnerability to popular revolution, in the absence of state breakdown at the top and

<sup>18</sup> Applied to the United States, d'Encausse's model would imply that the United States too, with its increasing disparity in ethnic birthrates, is heading toward disintegration along ethnic lines. Geopolitical theory, applied to the current conditions that predict high power-prestige for the United States in the middle-range future, predicts no such ethnic disintegration will occur.

<sup>19</sup> For the same kind of reasons, the prediction of downfall made by the Soviet dissident Andrei Amalrik (1970) must be taken as an ad hoc claim, since it was not based on principles of widely valid applicability.



conditions for resource mobilization from below (see evidence cited in Tilly [1978] and Skocpol [1979]).

*Capitalism versus socialism.*—Theory *b* should make us wary, since it is obviously a form of ideological gloating; but it has a certain plausibility, even for geopolitical theory, since if it is correct it would heavily influence the relative resource levels feeding into the military power of various states (i.e., it would be a noncumulative aspect of principle 1). To guard against post hoc imposition of theory upon a selected case, we should consider what general principles are being assumed. One is that economic productivity in consumer goods is the prime determinant of political revolutions. As a comparative-historical generalization this claim is plainly untrue.

A second weakness is that the theory of capitalism's productive superiority to socialism is vague with respect to timing. Its current application to the Soviet downturn of the 1980s is narrow and ad hoc; in an earlier period (through the 1950s and 1960s, until about 1975), Soviet socialist economies underwent more rapid growth than most of their capitalist counterparts (Kennedy 1987, pp. 429–31, 490–96). Szelenyi and Szelenyi (1994) argue on these grounds that the socialist economies may only have been in a medium-run cyclical phase of downturn in the 1980s and that the political downfall was due to factors that are fortuitous from the point of view of an economically driven theory. There is little to conclude from this argument either way, since there is no well-developed body of theory of long-term patterns of socialist economic change. Unfortunately, the same must be said about long-term patterns of capitalist economies. There is considerable evidence of cyclical processes and crises, but there are no general principles to indicate when and why capitalist economies might be in a cyclical crisis while socialist economies are growing (as in the 1930s) or when and why the reverse situation might exist (conceivably the case in the 1980s). In the future, social science may improve its understanding, and its predictive capabilities, regarding such macroeconomic patterns. But the balance of the evidence, and the coherence of theory development, suggests that even when we understand economic cycles better, they will not prove to independently determine state expansion or state breakdown; at best, they will add additional causal loops feeding into the core model of state resource extraction and its strains.

We also have available a crucial empirical comparison allowing us to test the relative merits of geopolitical theory and capitalist productive superiority theory: temporal patterns in state power of the Soviet Union and China. This evidence tells against the capitalist superiority model of state collapse. Li (1993) shows that the vicissitudes of Chinese Communist forces in the 20th century are those that follow from geopolitical

theory and predicts, on the basis of China's relative geopolitical advantages at the turn of the 21st century, that internal revolution is not to be expected. In other words, socialism *per se* does not account for variations in state weakness or strength, whereas geopolitical conditions do.<sup>20</sup>

The foregoing does not depend upon denying economic weakness within socialism but only whether it is possible in terms of general theory to derive from this fact an adequate explanation of the Soviet breakdown. There exists a well-developed body of analysis of the inefficiencies of a centrally planned economy of state enterprises (Kornai 1992). Moreover, as Walder (1994) and Nee and Lian (1994) show, attempting to reform such a system by introducing market structures weakens and splits the ruling elite by undermining the incentive structure for officials, decreasing their dependence upon central hierarchic controls, and promoting opportunistic seeking of personal economic gain. These processes cannot be the prime explanation of communist regime collapse, however; market reforms were little developed in the Soviet Union, the linchpin of Communist control in Eastern Europe, and most advanced in China of the 1980s, where the state remained strongest. Hungary and Poland, cases where extensive market reforms occurred, illustrate the meshing of the two kinds of processes. Market reform generated internal pressures for political change, but these led to full-scale state revolution only in the context of the Soviet Union's geopolitical crisis, which broke down control from above.

It may well be true, as Walder and Nee and Lian argue, that market reform will eventually undermine socialism in China. This market reform argument implies neither a state breakdown nor a democratic revolution. Indeed, insofar as economic growth through market reform further strengthens China's geopolitical position, one would expect a strong state, even as it evolves into some form of political economy other than pure socialism.<sup>21</sup>

<sup>20</sup> It should be noted that Goldstone (1992), applying his state breakdown model with emphasis upon the causal chain initiated by population growth, concludes that China in the late 20th century is in a prerevolutionary phase. We thus have a direct confrontation between the predictions of geopolitical theory and the population-centered variant of state breakdown theory. Future events will give further grounds for assessing the validity of these theories.

<sup>21</sup> A variant on the capitalist-superiority model is a stage theory of development, in which all other kinds of economies eventually give way to capitalism. Such theorizing may well be guided by the ideological euphoria of the period of anticommunist revolt. Empirically, there is as yet no clear evidence that a transition from socialism to capitalism is under way. As of the mid-1990s, it is not clear that many of the former Soviet bloc states are becoming full-fledged market capitalist economies; it is possible that hybrid or other forms will prevail, perhaps for long periods of time. A predominant form deriving from the disintegration of socialism seems to be industrial enterprise feudalism (Burawoy and Krotov 1992). The development of a well-supported

*Personality.*—Another factor retrospectively put forward to account for the Soviet overthrow is the influence of individual personalities. The underlying argument is an antitheoretical one: unique characters such as Gorbachev, springing up unpredictably, can make crucial turning points in world history. This line of analysis is easily disposed of. Such world-historical personalities do not appear at random; they have the possibility of world-historical significance only if they are structurally located in a position where their actions have major consequences. Such positions exist only when there are highly centralized structures of power and when power coalitions around that point are fragmented and in flux. In short, the conditions for the emergence of socially significant personalities are those formulated in state breakdown theory and its cognates in the theories of state formation and social movement mobilization. The sociological way to put the question is to ask for the conditions for the rise and fall of individual charisma.

The Gorbachev reform movement gives a particularly clear example of how charisma is dependent upon surrounding social dynamics. Gorbachev was an obscure figure, connected through Andropov with the secret police, until he became Communist Party chief in 1985. It was when he embarked upon the reform path, in response to structural crisis, that he became a charismatic figure; we can trace this new public personality to the mass meetings and tours that surrounded him with emotional fervor and made him an emblem for liberalizing (and peaceful) hopes.<sup>22</sup> This halo of charisma began to crumble as the crowds he had mobilized went beyond him and the coercive power of his regime was undermined, leaving him representing a weakening state authority. Great personalities rise and fall. The more general sociological lesson is to attend to the conditions under which we can expect such personalities to brighten and dim.

*Ideology.*—The Soviet collapse is sometimes cited as an instance of the overriding role of ideology. The downfall is variously attributed to the irrepressible idea of ethnic nationalism, the spreading ideas of freedom, or the idea of the capitalist market (more specifically, its consumerist ethic). There is no reason to doubt the existence of these beliefs. The questions are, Did they determine the Soviet transition? and Does ideology provide a basis for prediction generally, or for concluding that prediction is impossible?

---

theory of economic transitions still has a long way to go; in its absence, we have little to rely on here for systematic predictions.

<sup>22</sup> Thus Gorbachev's visit to China in May 1989, in an effort to reduce geopolitical confrontation, was the catalyst for the mass demonstration at Tiananmen Square. The failure of that uprising shows that individual charisma alone is insufficient to produce structural change in the absence of the factors listed in state breakdown theory (see Li 1993).

In the case of ethnic ideologies, structural conditions for this type of mobilization have been discussed above. Ethnic nationalism always potentially exists, but whether it goes in the direction of assimilation to larger national units (a dominant trend in recent centuries; see Tilly 1993, pp. 246–47) or toward local particularism is highly variable. It is tautological to conclude that the autonomous strength of ethnic ideology *per se* determines which way it will go. The same can be said about the ideological appeals of capitalism. Procapitalist ideology does not have a uniform appeal, even on a decade-by-decade basis in the 20th century, and there have been a number of reversals in the popularity of pro- and anticommunist ideologies. When and where these variations occur has not been systematically theorized. If we limit the capitalist ideology thesis to the spread of Western consumerist attitudes, it is striking that the contemporary ideology among Western intellectuals and mass-entertainment culture was, far from celebrating consumerism, a post-modernist attack upon it. Thus the ideology of the antisocialist dissidents was out of phase with the alleged transmitters of ideology in the West.

More theoretically, the role of ideology in state breakdowns and revolutions has been addressed by the cumulating body of research that I have drawn upon here. Ideologies associated with the directions that revolutions take have been quite variable, and theoretical explanation has not been well developed. Goldstone (1992, pp. 416–48) argues, like Skocpol, that ideologies are not a significant part of the explanation of state breakdowns. (In other words, ideology is a proximal, not a basic, part of the causal chain; it comes into play only once inraelite conflict and state fiscal crisis are well advanced. Compare the formulation of Moaddel [1992].) For Goldstone, ideology is a better predictor of what follows once breakdown takes place. Nevertheless, we lack systematic theory of what ideology that will be for particular kinds of cases. There is considerable evidence that ideological movements “march backward into the future.” For instance, the major outburst of class conflict in the early industrial revolution in England was mobilized under a reactionary ideology of workers attempting to turn back the clock to artisan production but succeeded only in laying the basis for the regulation of factory production (Calhoun 1982).

My Russian empire prediction included the theoretical argument that ideology follows geopolitics. For the parts of the Soviet Union that fell within the geopolitical orbit of nearby Islamic states, I proposed that the spread of Islam would be the vehicle for revolt. The prestige of an ideology rises and falls with the state power of its most visible adherents; thus the success of the Iranian revolution under Islamic fundamentalist auspices led to the rapid upsurge of Islamic fundamentalism in adjacent

areas, including Afghanistan. For the bulk of the Soviet Union, I suggested that state breakdown and regime transition might well occur under the ideology of a dissident form of Communism. This turned out to be how it got started (Gorbachev's reform Communism), but I did not anticipate that the breakdown would rapidly turn in a procapitalist direction. Here I might have made better use of the general principle that ideologies, like rulers, are delegitimated by falls in the power-prestige of the regime that carries them. Just as defeat in World War I undermined the ideological prestige of the capitalist modernizers who controlled the Russian government and set the stage for a rebound into anticapitalist ideology, the geopolitically based failure of the Soviet Union (and the continuing downward slope of power-prestige for Gorbachev's reform Communists) led to a rebound to anticommunist and procapitalist ideology.<sup>23</sup>

State breakdowns thus would tend to bring about shifts from one of a pair of ideological rivals to another. Since ideologies are not usually limited to a pair of alternatives, the generality of this principle does not go far. We need better theorizing of the conditions that produce ideologies and elevate one or another of them to prominence.<sup>24</sup> Ideology can no doubt be incorporated into a general, predictive theory of state breakdowns, but instead of making it a free-floating unmoved mover, we need to establish the conditions for its variants and the place they hold in the causal chain.

In sum, the rival explanations of Soviet breakdown reviewed here have been *ad hoc*. They do not rest on general principles consistent with explanation across a broad range of cases. Geopolitical theory, in contrast, has been based on comparative data from the outset. Geopolitical

<sup>23</sup> The ideological prestige of capitalism throughout its existence may thus be connected with its geopolitical associations as much as its economic performance. It is consistent with this argument that the ideology of capitalism spread around the world in conjunction with British and then U.S. geopolitical strength and came under attack at just the times and places where geopolitical control weakened. The issue demands more systematic investigation.

<sup>24</sup> Wuthnow's (1989) systematic comparisons of which European states adopted or rejected Protestantism, the Enlightenment, and the 19th-century socialist movements provides perhaps our most advanced model here. Wuthnow's state-centered theory bears a kinship to the Skocpol-Goldstone-Tilly model of state resource extraction. A central process behind major new ideological movements is the emergence of new economic and organizational resources that strengthen the state sector. When this happens in conjunction with a stalemate between state actors and the property-controlling conservatives (a variant on the Skocpol model), opportunity is opened up for a third faction of cultural entrepreneurs; their success depends upon the degree of stalemate between the two branches of the intraelite conflict, together with the extent of expansion of new material and organizational means of cultural production.

theory of state change is still far from having a high level of refinement or of having gone through stringent empirical testing. Nevertheless, it makes roughly accurate predictions and explanations. It meshes well with the major lines of theory and research on the development of the state and on state breakdowns and revolutions. The details of how the Soviet breakdown actually took place are consistent with the processes expected from these models. It is reasonable to conclude that geopolitical theory is a building block from which further prediction-generating work can be mounted (see appendix).

### How Much Predictive Precision Is Possible?

Current geopolitical theory is not precise. Historical atlases show that the conditions that determine geopolitical advantages and disadvantages change rather slowly, over periods of hundreds of years (Collins 1978). Since the state controls military force over a territory, the relative lineup of such resources pays off in much shorter periods, during times of wars and related acute periods of resource strains; major wars typically take from two to five years, rarely more than 10. This gives us two different time orders: the long slow latent buildup or build-down of resources, punctuated by war times when such shifts are manifested in overt changes in territorial power. The conditions that determine when wars actually break out, or correspondingly when wars are avoided, include a great many other factors.<sup>25</sup> These factors are apparently randomized over the long run, however, as we can infer because the long-term direction of relative resources eventually leads to corresponding state expansion or contraction. When we apply geopolitical theory to actual predictions, then, we are left with some inherent imprecision. From historical atlases, I estimated that geopolitical resources give predictability down to units of about 30–50 years; within such periods, it is impossible to know (from geopolitical information alone) when the crucial military-driven crises will occur. From the standpoint of 1980, I predicted that the Soviet Union would disintegrate within 30–50 years. Frankly, I was surprised that it happened so soon, but it certainly was within the scope of my prediction.<sup>26</sup>

<sup>25</sup> There has been some success in predicting war outbreaks, given the existence of particular starting points at which international disputes occur; an empirically successful model builds upon the relative resources and alliance patterns of the opposing states (Bueno de Mesquite and Lalman 1992).

<sup>26</sup> Not every aspect of the Soviet breakdown was included in my geopolitical prediction. It has been pointed out that the transition was relatively bloodless, with prolonged revolutionary fighting only in one instance (Romania). The implication seems

Because of factors on this intermediate or mesolevel of causality—for instance, because diplomatic alliances that offset any particular strength or weakness can always be made—it has sometimes been argued, much more sweepingly, that geopolitical resources are inherently incapable of generating predictions of state power changes. Mann (1989) argues against Kennedy's (1987) formulation of geopolitical theory, saying that any such model is invalid because diplomatic alliances are unpredictable. But is it in fact the case that diplomacy is a realm of free-floating choices? There has been as yet no systematic confrontation of the literature on diplomacy with that of geopolitical theory, but it is reasonable to expect that diplomacy follows geopolitics. I propose two hypotheses:

**HYPOTHESIS 1.**—*Geopolitically strong states impose alliances upon weaker states adjacent to their immediate zone of military extension; conversely, weak states seek the protection of adjacent strong states or give in to the imposition of alliance.*

**HYPOTHESIS 2.**—*Where balance-of-power situations obtain (i.e., in zones where multiple states impinge upon each others' borders), states make alliances on the principle of "my enemy's enemy is my friend."* These alliances lead to a geopolitical checkerboard pattern and the fragmented interior region predicted by geopolitical principle 3. These middle or interior regions exist only as long as marchland states are not strong; when power resources cumulatively build up on the opposite edges of such a middle region, alliances shift to bipolar blocs dictated by the largest states. Historical cases ranging from the Roman expansion to the Soviet and U.S. alliances of the 20th century suggest that, far from offsetting geopolitical principles, diplomacy is geopolitics by other means.

There is a third order of time, much shorter than either of the above. This is social movement time. At its heart are the two or three days during which the state hangs in the balance, the crowds are mobilized in the streets, and waves of enthusiasm or fear forge massive social networks, not least of which are the soldiers wavering from one center of power to another. Around this microfocus of time when the state power transition actually occurs—the revolutionary "days" of popular fame—there is a penumbra of semi-intense mobilization, lasting a few weeks. This period may be stretched out into months if movement enthusiasm

---

to be that this pattern invalidates a theory of revolutionary change, at any rate that theory embodied in Marx's adage "force is the mid-wife of every society pregnant with a new order." Geopolitical and state breakdown theory, however, do not imply that every transition is violent but only that the accumulation of resource strains based on the organized means of violence sets in motion processes that eventuate in regime change.

diffuses among a series of state power centers, such as the various capitals of Eastern Europe in the fall and winter of 1989–90 or the regional republics of the Soviet Union from 1990 onward.<sup>27</sup>

This microfocus of mobilization time is so emotionally intense and so capable of forging symbols that it tends to obscure the other two orders of time processes (decade- or century-long shifts in relative state resources and the several years of military war strains) that made it possible. It was a journalistic cliché, especially in the summer of 1991, to remark how amazing it was that change could occur in the Soviet regime so rapidly. It would have been sociologically amazing, however, if the change in state control had *not* occurred rapidly. At the core of the state (the monopoly of organized military force) is a coercive coalition, enforcers who discipline one another by armed threat. When such a coalition breaks apart and is replaced by another coalition, it must go through a rapid tipping point, since it is extremely dangerous for an individual enforcer to stay outside the winning coalition. These tipping points generate extreme and contagious emotions, and they give the sense that everything hinges in the balance, that there is a moment of freedom and choice. This is what generates the ideology that revolutions are uncausable and unpredictable.

Nevertheless, the passage through such tipping points is itself highly structured by the two prior orders of time processes that I have discussed. Just when this tipping point comes to pass is unpredictable, at least from a more macro standpoint, but that such tipping points will occur within a larger time frame is predictable when the macro-orders of causality have moved beyond a certain degree.

*Nested levels of macro- to microlevel prediction.*—What I have been considering here is a macro/micro problem. The issue of the different orders of causal and predictive precision points up nicely the character of the relationship between the macrolevel and microlevel. What we are concerned with is a continuum, not a dichotomy. The macrolevel end contains those patterns of causal connection that are relatively biggest in time and space, while the more microlevel end includes the patterns of social organization that can be discerned in successively smaller slices of time and space. Thus the relationship between relatively more macro and

<sup>27</sup> These recent cases appear to be part of a more general process of revolt contagion in interlinked states; a prime example is the chain of revolts in the spring and summer of 1848, spreading from Switzerland, Italy, and France through the German states, Austria, and Hungary, then demobilizing during the next year. Another such wave of revolts happened in cities throughout Germany and central Europe at the close of World War I (Kinder and Hilgemann 1968, pp. 328, 406). A theory of revolt contagion has yet to be formulated.



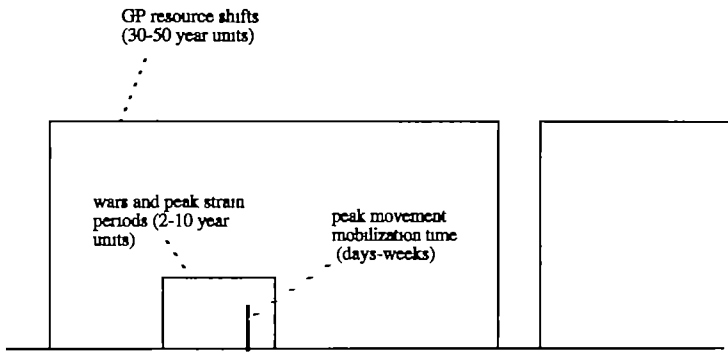


FIG. 6.—Nested macro-micro levels of predictability in state breakdown

more micro is one of nested levels (see fig. 6). The geopolitical macrolevel consists in causal patterns that cover periods of decades up through centuries, as well as variability in the degree of connectedness among the social organizations in different regions of space. Its predictive generalizations can at best issue in statements about the direction of change in geopolitical advantages and disadvantages and point to shifts that take place sometime within a range of 30–50 years.

Nested inside such time periods are moderately large-scale processes such as wars and state breakdowns. (Given the imprecision of our vocabulary for such matters, we might refer to them as *lower macro* or *upper meso*; they link thousands or millions of mutually unseen actors into chains of relationships that may unfold over a period of years.) The problem of predictive imprecision on this upper mesolevel applies primarily to the initiation of a conflict or breakdown, which may be inherently random insofar as it depends upon the concatenation of a number of smaller crises, as in Perrow's (1984) "normal accidents." However, once a breakdown process is under way, it displays a great deal of *theoretical* order, in the sense that abstract properties of the unraveling of coercive coalitions and their replacement by others are widely, even universally, found. Recall, however, that successful prediction requires a combination of valid general theory plus empirical knowledge of relevant starting points; prediction is often not feasible in the heat of unfolding events, because information on these empirical conditions is not available.

Focusing in still further, we reach the lower mesolevel of movement mobilization time and, nested within it, the heady hours of peak collective behavior time.<sup>28</sup> Implicitly, these lower mesolevel events that we

<sup>28</sup> I refer to this as *lower meso*, reserving *micro* for the truly small slices of interaction between a few face-to-face actors, down to the ultramicrolevel of conversational and emotional or nonverbal rhythms.

consider are chosen from a wider selection of similar time periods in which routine prevails and apparently nothing worth analyzing happens. Once again, the crucial predictive imprecision has to do with the timing of when a particular kind of activity in this region in space begins. Kuran's (1995) companion paper in this symposium gives a stronger reason why the exact timing of these tipping points is socially unpredictable by the participants themselves. This reason is a theory of unpredictability on the third level of the nested macro-micro continuum, dealing with the initiation point for the outbreak of oppositional social movements. The misrepresentation of preferences, which is the normal condition of persons within a mutually enforced coercive coalition, can itself only break down under special conditions (which ultimately derive from the nested macrolevel of processes on the state breakdown level). Once the tipping point is passed, however, the social movement dynamics and the peak periods of collective behavior follow highly structured patterns. Crowds, even at the height of emotional enthusiasm, do not behave randomly but exhibit considerable social coordination down to the level of true micropatterns (McPhail 1991, pp. 158-84). Such tipping points may also be characterized as passage through conditions of critical mass; this passage occurs under network conditions specified by Marwell and Oliver (1993).

In short, the problem of predictive imprecision occurs at just those points where we shift from a larger level of macroanalysis to a different level of social interaction nested within it; the nesting within refers us directly to a shift in time orders, which is what the imprecision is about. This is a way of saying that, from considerations of factors that work themselves out over a given stretch of time (e.g., half-centuries), it is impossible to predict with any greater precision the onset of processes that work themselves out in 3-10 years; the same is true in the relation among successively narrower nestings in time.

### Obstacles to Successful Sociological Prediction

There are a number of reasons why sociologists do not make more good predictions. Briefly, these include the failure to distinguish empirical extrapolation from theory-based prediction and the failure of theorists to collect adequate data about empirical starting points from which predictions may proceed. In addition, there is a considerable variety of competing theories in many areas and relatively little focus of attention on which lines of theory have the highest level of validation (by the criteria that I have discussed above). Moreover, a great deal of metatheoretical discussion in sociology is concerned not with any particular substantive explanation but with abstract reasons why substantive explanation (and a

fortiori prediction) of human action is impossible and perhaps also immoral. Without entering into lengthy debates over historicism, interpretivism, and humanism, I would only note that the impossibility of prediction is refuted in the present case and that there are numerous other areas in sociology where an appropriate focus upon the best-validated explanatory theory, together with sufficient information of empirical starting points, can yield other successful predictions.<sup>29</sup>

The foregoing obstacles to prediction are internal to the community of social scientists. There are also external conditions that are especially likely to be obstacles to macrosociological prediction of politically significant events. One such influence comes from the political ideologies of the surrounding society, in which we are also participants. I have already noted that, through most of the 1980s, both liberals and conservatives were locked into a stereotyped image of Soviet power. By the late 1980s, Gorbachev's liberalizing reforms were well under way, but U.S. newspapers relegated them to the inside pages. During the 1988 presidential campaign, the challenger, Michael Dukakis, ignored foreign policy initiatives toward military deescalation, even though polls showed that Gorbachev was more popular in the United States than any American politician. Through the early summer of 1989, only months before the overt breakdown began in Eastern Europe, President Bush and his administration took the line that the reforms were only a ruse and that continued military buildup was necessary (*Facts on File World Political Almanac* 1986–89).

Can we extract a sociological lesson from the behavior of American politicians during the late 1980s about the dynamics of ideology? It appears that ideologies, forged during previous conflicts, are slow to change even when underlying structural conditions change. It is only when there is a highly dramatized, overt change in surface relationships that a new set of ideological symbols comes to the fore. Before 1989 the dominant perspectives worldwide saw the communist states as permanent and powerful; within a few years, it became equally common belief that they had been doomed to failure. Neither view was (or is) based upon well-grounded theoretical understanding of the bases of state power. Social scientists are capable of contributing to such understanding to the degree that they can insulate themselves, at least in a portion of their lives, from popular political ideologies.

That this is difficult to do is illustrated by the case of Paul Kennedy. Kennedy published in 1987 a general treatment of rises and declines in

<sup>29</sup> The issue of the inhumanness or immorality of predictions must be discussed elsewhere. In the present case, I feel there was nothing immoral about attempting to make a contribution in 1980 to surviving the nuclear arms race.

state power, with an application to all the major states of the contemporary world. Kennedy independently formulated several of the major principles of geopolitical theory that I have listed above, basing his analysis upon the principles of resource advantage (1) and overextension (4). Nevertheless, he missed the application of these principles to the future decline of Soviet power, because his main concern was the danger of U.S. decline, above all as a result of overextension of U.S. military power around the globe. The influence of ideological commitment is visible here. Kennedy expressed the viewpoint of American liberals, concerned to avoid repeating the mistake of the Vietnam war; his theory encompassed this concern under the principle of overextension of military power to distant regions, resulting in power decline by attrition of resources. Kennedy's theoretical principles, and the historical comparisons behind them, are correct as far as they go; his prediction failed because ideological commitments, which focused his attention on the state that he most wished to warn—his own—while causing him to lose perspective regarding the one in which geopolitical strains were greatest.

There are also more routine reasons why we do not develop or make use of the predictive power of explanatory sociology. Ordinary discourse and the rituals of social interaction tend to reify social institutions, so that they appear permanent and immutable. Ethnomethodological investigations of everyday reasoning show that social actors prefer to take for granted a constant background of normalcy; when some incident forces a breach in normal expectations, we try to mend the breach by reimposing a normal form upon it as soon as possible. This is one reason why ordinary social actors perceive the macrostructures of the world through the lens of ideologies, even though these ideologies do little more than impose an arbitrary sense of order upon the world. Nevertheless, the routine of sociological researchers and theorists is not the same as the routine of lay actors; we have built up intellectual resources that enable us to piece together coherent and empirically validated models of social processes, whose categories need not be the same as those of the taken-for-granted categories of everyday life. It is demonstrable that sociological theory may be used for valid predictions that go beyond the "unaided eye." This use of theory makes it possible for sociology to make a distinctive contribution.

#### Future Prospects for Predictive Sociology

The ability of sociology to make valid predictions is a sign of the maturity of the discipline. Valid prediction is not a one-shot enterprise; it rests upon the cumulative development of theory and research. Investigating promising leads and blind alleys, accumulating data and discovering dis-

tinctions and interrelations, grasping complexity and strategically simplifying core models—all this takes generations. It would be surprising if no such progress had been made during the 100 years of sociology's institutional existence, roughly since the founding of the *American Journal of Sociology*, contemporary with the academic and professional organization of sociology in both the United States and France. The macrodynamics of political change is one of the longest standing research interests in sociology; the passion and energy it has attracted over the years has given it a core of theory that can provide increasingly good service as more refined theory is elaborated in the future.

## APPENDIX

### Further Applicability of Geopolitical and State Breakdown Theory

To illustrate the general applicability of geopolitical and state breakdown theory, consider briefly the assessment of future political stability in two trouble spots of the mid-1990s. South Africa has strong geopolitical advantages vis-à-vis its neighbors on the African continent (and in relation to several tiers of states to the north): it has size and resource advantages of the largest population and economy; it occupies the marchland position against multifront states; there are no other strong and potentially expansive states remotely near it; in its present condition there is no danger of military overextension, although overextension might become a problem many decades in the future if South Africa makes use of its geopolitical advantages to establish an empire of client states. The main condition until 1994 preventing South Africa from expansion has been the internal warfare between black and white populations; in an earlier publication (Collins 1978, p. 32) I suggested that when the black majority took political control, South Africa would go into an expansive phase in the following decades. Insofar as external conflict brings solidarity and legitimacy follows from interstate power-prestige, the expansion of South African military power would favor the successful institutionalization of its regime. (Note that expansion is not necessarily based on overt conquest; it could take a form similar to U.S. foreign power in the post-1945 period: leading coalitions of allies and exercising peacekeeping missions.)

A contrasting case is the geopolitical prediction for the future of the former Yugoslav federation and its immediate neighbors. Here mixtures of population and economic resources prevent the clear advantage for any of these states and state fragments; all are in multifront situations of one degree or another. (Hence, historically, they have been unified only by large empires based in the more concentrated resource zones to the northwest or southeast.) This is the formula for shifting balance-of-power

politics; the ethnic checkerboard is the result of many centuries of such geopolitical changes. Until such time as the European Community becomes an integrated and expansionary military power, taking client states under its wings, one can expect that this region of the Balkans will be militarily volatile, with low regime legitimacy.

## REFERENCES

- Amalrik, Andrei. 1970. *Will the Soviet Union Survive until 1984?* New York: Harper & Row.
- Andreski, Stanislaw. 1971. *Military Organization and Society*. Berkeley: University of California Press.
- Becker, Abraham S. 1986. *Sitting on Bayonets: The Soviet Defense Budget and the Slowdown of Soviet Defense Spending*. Santa Monica, Calif.: RAND/UCLA Center for the Study of Soviet International Behavior.
- . 1987. *Ogarkov's Complaint and Gorbachev's Dilemma: The Soviet Defense Budget and Party-Military Conflict*. Santa Monica, Calif.: RAND.
- Bendix, Reinhard. 1967. "Tradition and Modernity Reconsidered." *Comparative Studies in Society and History* 9:292–346.
- Bernstein, Alvin H. 1989. *Soviet Defense Spending*. Santa Monica, Calif.: RAND.
- Bessinger, Mark R. 1990. *Nonviolent Public Protest in the USSR, December 1, 1986–December 31, 1989*. Washington, D.C.: National Council for Soviet and East European Research.
- Boulding, Kenneth. 1962. *Conflict and Defense*. New York: Harper & Row.
- Bueno de Mesquite, Bruce, and David Lalman. 1992. *War and Reason*. New Haven, Conn.: Yale University Press.
- Bueno de Mesquite, Bruce, Randolph M. Siverson, and Gary Woller. 1992. "War and the Fate of Regimes: A Comparative Analysis." *American Political Science Review* 86:638–46.
- Burawoy, Michael, and Pavel Krotov. 1992. "The Soviet Transition from Socialism to Capitalism." *American Sociological Review* 57:16–38.
- Calhoun, Craig. 1982. *The Question of Class Struggle*. Chicago: University of Chicago Press.
- Chase-Dunn, Christopher. 1989. *Global Formation: Structures of the World Economy*. Oxford: Blackwell.
- Collins, Randall. 1978. "Long-Term Social Change and the Territorial Power of States." *Research in Social Movements, Conflicts, and Change* 1:1–34.
- . 1981. "Does Modern Technology Change the Rules of Geopolitics?" *Journal of Political and Military Sociology* 9:163–77.
- . 1986. *Weberian Sociological Theory*. New York: Cambridge University Press.
- . 1993. "Maturation of the State-Centered Theory of Revolution and Ideology." *Sociological Theory* 11 (March): 117–28.
- Collins, Randall, and David V. Waller. 1992. "What Theories Predicted the State Breakdowns and Revolutions of the Soviet Bloc?" *Research in Social Movements, Conflicts, and Change* 14:31–47.
- . 1994. "The Geopolitics of Ethnic Mobilization: Some Theoretical Projections for the Old Soviet Bloc." Pp. 79–104 in *Legacies of the Collapse of Marxism*, edited by John H. Moore. Arlington, Va.: George Mason University Press.
- Coser, Lewis A. 1956. *The Functions of Social Conflict*. Glencoe, Ill.: Free Press.
- Downing, Brian M. 1993. *The Military Revolution and Political Change: Origins of*

- Democracy and Autocracy in Early Modern Europe*. Princeton, N.J.: Princeton University Press.
- d'Encausse, Helene C. 1979. *Decline of an Empire: The Soviet Socialist Republics in Revolt*. New York: Harper & Row.
- Enggass, P. M. 1986. *Geopolitics: A Bibliography of Applied Political Geography*. Monticello, Ill.: Vance Bibliographies.
- Evans, Peter B., Dietrich Rueschemeyer, and Theda Skocpol, eds. 1985. *Bringing the State Back In*. New York: Cambridge University Press.
- Facts on File World Political Almanac*. 1984-91. New York: Facts on File.
- Gaddis, John Lewis. 1992. "International Relations Theory and the End of the Cold War." *International Security* 17 (3): 5-58.
- Gelman, Harry. 1989. *The Soviet Turn toward Conventional Force Reduction: The Internal Struggle and the Variables at Play*. Santa Monica, Calif.: RAND.
- Gilpin, Robert. 1981. *War and Change in World Politics*. New York: Cambridge University Press.
- Goldstone, Jack A. 1991. *Revolution and Rebellion in the Modern World*. Berkeley and Los Angeles: University of California Press.
- . 1992. "Immanent Political Conflicts Arising from China's Environmental Crisis." Occasional Papers Series of the Project on Environmental Change and Acute Conflict. American Academy of Arts and Sciences, Cambridge, Mass.
- Hanneman, Robert A., Randall Collins, and Gabriele Mordt. 1995. "Discovering the Long-Term Dynamics in Abstract Theory: Simulation Experiments with State Legitimacy and Imperialist Capitalism." *Sociological Methodology*, vol. 25. In press.
- Hepple, Leslie W. 1986. "The Revival of Geopolitics." *Political Geography Quarterly* 4 (4): S21-S36.
- Keating's Record of World Events*. 1984-91. London: Longman.
- Kennedy, Paul. 1987. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House.
- Kinder, Hermann, and Werner Hilgemann. 1968. *Atlas Historique*. Paris: Stock.
- Kornai, Janos. 1992. *The Socialist System: The Political Economy of Communism*. Princeton, N.J.: Princeton University Press.
- Kuran, Timur. 1995. "The Inevitability of Future Revolutionary Surprises." *American Journal of Sociology*, in this issue.
- Li, Jie-li. 1993. "Geopolitics of the Chinese Communist Party in the Twentieth Century." *Sociological Perspectives* 36:315-33.
- Liddell-Hart, B. H. 1970. *History of the Second World War*. New York: Putnam.
- Luttwak, Edward. 1976. *The Grand Strategy of the Roman Empire: From the First Century A.D. to the Third*. Baltimore: Johns Hopkins University Press.
- Mann, Michael. 1986. *The Sources of Social Power*, vol. 1. New York: Cambridge University Press.
- . 1989. "Comments on Paul Kennedy's *The Rise and Fall of the Great Powers*." *British Journal of Sociology* 40:331-35.
- Marwell, Gerald, and Pamela Oliver. 1993. *The Critical Mass in Collective Action: A Micro-Social Theory*. New York: Cambridge University Press.
- McNeill, William H. 1963. *The Rise of the West: A History of the Human Community*. Chicago: University of Chicago Press.
- McPhail, Clark. 1991. *The Myth of the Madding Crowd*. New York: Aldine.
- Moaddel, Mansoor. 1992. "Ideology as Episodic Discourse: The Case of the Iranian Revolution." *American Sociological Review* 57:357-79.
- Modelska, George, and W. R. Thompson. 1988. *Sea Power and Global Politics since 1494*. Seattle: University of Washington Press.

## American Journal of Sociology

- Morgenthau, Hans J. 1948. *Politics among Nations*. New York: Knopf.
- Moore, Barrington, Jr. 1966. *Social Origins of Dictatorship and Democracy*. Boston: Beacon.
- Nee, Victor, and Peng Lian. 1994. "Sleeping with the Enemy: A Dynamic Model of Declining Political Commitment in State Socialism." *Theory and Society* 23: 253-96.
- Norpoth, Helmut. 1987. "Guns and Butter and Government Popularity in Britain." *American Political Science Review* 81:949-70.
- Office of Technology Assessment, U.S. Congress. 1979. *The Effects of Nuclear War*. Washington, D.C.: Government Printing Office.
- Ostrum, Charles W., Jr., and Dennis M. Simon. 1985. "Promise and Performance: A Dynamic Model of Presidential Popularity." *American Political Science Review* 79:334-58.
- Paige, Jeffrey M. 1975. *Agrarian Revolution*. Berkeley: University of California Press.
- Parker, Geoffrey. 1988. *The Military Revolution: Military Innovation and the Rise of the West*. New York: Cambridge University Press.
- Perrow, Charles. 1984. *Normal Accidents*. New York: Basic.
- Roeder, Philip G. 1991. "Soviet Federalism and Ethnic Mobilization." *World Politics* 43:196-232.
- Schelling, Thomas C. 1962. *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.
- Sedaitis, Judith B., and Jim Butterfield, eds. 1991. *Perestroika from Below: Social Movements in the Soviet Union*. Boulder, Colo.: Westview.
- Skocpol, Theda. 1979. *States and Social Revolutions*. New York: Cambridge University Press.
- Singer, J. David. 1979. *Explaining War*. Beverly Hills, Calif.: Sage.
- Singer, J. David, and Paul F. Diehl. 1990. *Measuring the Correlates of War*. Ann Arbor: University of Michigan Press.
- Stinchcombe, Arthur L. 1961. "Agricultural Enterprise and Rural Class Relations." *American Journal of Sociology* 67:165-76.
- . 1968. *Constructing Social Theories*. New York: Harcourt Brace.
- Stone, Lawrence. 1979. *The Family, Sex and Marriage in England, 1500-1800*. London: Weidenfeld & Nicolson.
- Szelenyi, Ivan, and Balazs Szelenyi. 1994. "Why Socialism Failed: Toward a Theory of System Breakdown—Causes of Disintegration of East European State Socialism." *Theory and Society* 23:211-31.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.
- . 1990. *Coercion, Capital, and European States, A.D. 990-1990*. Oxford: Blackwell.
- . 1993. *European Revolutions, 1492-1992*. Oxford: Blackwell.
- Thompson, William R. 1988. *On Global War: Historical-Structural Approaches to World Politics*. Columbia: University of South Carolina Press.
- Van Creveld, Martin. 1977. *Supplying War: Logistics from Wallenstein to Patton*. New York: Cambridge University Press.
- . 1991. *Technology and War: From 2000 B.C. to the Present*. New York: Free Press.
- Walder, Andrew G. 1994. "The Decline of Communist Power. Elements of a Theory of Institutional Change." *Theory and Society* 23:297-324.
- Waller, David V. 1992. "Ethnic Mobilization and Geopolitics in the Soviet Union: Toward a Theoretical Understanding." *Journal of Political and Military Sociology* 20:37-62.
- Wallerstein, Immanuel. 1974-89. *The Modern World System*, vols. 1-3. San Diego,



- Calif.: Academic Press.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: Columbia University Press.
- Weber, Max. [1922] 1968. *Economy and Society*. New York: Bedminster Press.
- Wuthnow, Robert. 1989. *Communities of Discourse: Ideology and Social Structure in the Reformation, the Enlightenment, and European Socialism*. Cambridge, Mass.: Harvard University Press.

# To Explain Political Processes<sup>1</sup>

Charles Tilly

*New School for Social Research*

Analysts of large-scale political processes frequently invoke invariant models that feature self-contained and self-motivating social units. Few actual political processes conform to such models. Revolutions provide an important example of such reasoning and of its pitfalls. Better models rest on plausible ontologies, specify fields of variation for the phenomena in question, reconstruct causal sequences, and concentrate explanation on links within those sequences.

Asked to explain particular instances of vigilante violence, social movements, citizenship, wars, nationalism, or transformation of states, sociologists search almost instinctively for general, invariant models of those phenomena to which they can assimilate the cases at hand. Reflecting on why the disintegration of the Soviet Union and its Warsaw Pact allies so surprised Western analysts, sociologists immediately wonder: Into what general category of recurrent events should they have put the Eastern European experience of the 1980s? Does it belong to revolution, nationalism, democratization, political modernization, imperial disintegration, or something quite different? Sociologists suppose that if they had recognized the category when the process began they would have been able to predict its outcome.

S. N. Eisenstadt (1992, p. 21), for example, places the breakdowns of communist regimes in parallel with other revolutions: "Are these revolutions 'great revolutions'—the English civil war, the American, French, Russian, and Chinese revolutions—which in many ways ushered in modernity and created the modern political order? Are they likely to lead—after a possibly turbulent period of transition—to a relatively stable world of modernity, with liberal constitutionalism heralding some kind of 'end of history'? Or do they tell us something of the vicissitudes and fragilities of modernity, even of democratic-constitutional regimes?"

<sup>1</sup> Correspondence may be addressed to Charles Tilly, Center for Studies of Social Change, New School for Social Research, 64 University Place, New York, New York 10003-4520.

Eisenstadt replies that some part of each of these questions is true. The Eastern European transitions qualify as revolutions, he says, but unlike their predecessors they set their faces against, rather than for, modernity. In the very process of identifying differences, Eisenstadt reinforces the idea that a useful model of revolution specifies similarities, invariant general processes. There he joins the majority of other analysts.

As Eisenstadt does, moreover, sociologists usually assume that the processes in question occur within self-contained social units—societies, states, aggrieved populations, or something of the sort—in a self-propelled way. They assume coherent, durable monads rather than contingent, transitory connections among socially constructed identities. We can hardly blame them for it; we veterans taught them to do it in graduate school, because that is also what we learned to do in graduate school years earlier. We learned and in turn taught a practice of this sort: (1) assume a coherent, durable, self-propelling social unit; (2) attribute a general condition or process to that unit; (3) invoke or invent an invariant model of that condition or process; (4) explain the behavior of the unit on the basis of its conformity to that invariant model.

The most egregious examples of invariant thinking appear in comparative-historical analyses where nations, states, or societies serve as the objects of comparison. Even methodological individualists frequently follow the same logic, albeit on a smaller scale. They model the necessary/sufficient conditions under which a rational decision maker (or, in other versions, the follower of a unitary vision, illusion, or impulse) would take steps to create a state, start a war, rebel, secede, vote, join a social movement, or carry on some other well-defined political performance.

Similar reasoning appears frequently in studies of nationalism, democratization, the disintegration of empires, social movements, transformations of states, wars, revolutions, and other large-scale political phenomena. In the case of nationalism, available theories range from primordialist to constructivist, from realist to subjectivist, but a surprising proportion of them claim not to account for the variable degrees or qualities of nationalism but to place most or all nationalisms in the same box (for convenient surveys, see Anderson 1991; Comaroff and Stern 1993; Connor 1987; Feschbach 1987; Gellner 1983; Haas 1986; Hobsbawm 1990; Kearney 1991; Lerner 1991; Lowi 1992; Löwy 1989; Segal 1988; Williams 1989). The study of social movements offers more promising recent trends, since a number of scholars have taken to relating variation in the organization of movements systematically to differences and fluctuations in political opportunity structure (e.g., Duyvendak 1994; Giugni and Kriesi 1990; Koopmans 1993; Kriesi 1993; Tarrow 1993). Yet even in this area much theorizing has proceeded as if all social movements fell into just two internally homogeneous categories: old and new (Cohen

1985; Diani 1992; Eyerman and Jamison 1991; Mayer 1991; Melucci 1985, 1989).

In analyses of democratization and state formation, similar simplifications prevail (for surveys, see Alker 1992; Barkey and Parikh 1991; Bratton 1989; Caporaso 1989; Dahl 1989; Diamond and Marks 1992; Di Palma 1990; Gurr 1988; Gurr, Jagers, and Moore 1990; Hall and Ikenberry 1989; Held 1987; Kirby and Ward 1991; Krasner 1984; Lee 1988; Mann 1990; Mitchell 1991; O'Donnell and Schmitter 1986; Poggi 1990; Rueschemeyer, Stephens, and Stephens 1992; Schmitter and Karl 1991). Most and Starr have long since offered the same kind of complaint about studies of war (1983; see also Levy 1989; Starr 1994). In short, invariant models concerning self-motivating social units continue to wind like honeysuckle through the study of large-scale political processes.

If it is examined closely, the standard practice makes little sense. Coherent, durable, self-propelling social units—monads—occupy a great deal of political theory but none of political reality. Ostensible general conditions such as revolution, nationalism, or war always turn out to fall not at a single point but to stretch along a whole range of positions on some intersecting set of continua. The employment of invariant models, furthermore, assumes a political world in which whole structures and sequences repeat themselves time after time in essentially the same form. That would be a convenient world for theorists, but it does not exist.

Although the assumption of sharply bounded, self-motivating social units deserves equal criticism, William H. Sewell, Jr. (1992), Margaret Somers (1992), Harrison White (1992), and others have recently criticized monadic thinking so effectively that—however much I disagree with some of their proposed remedies—I have little to add to their critiques. Let me therefore concentrate here on the assumption of invariant conditions and processes. The general structure runs like this:

1. All *A*'s have characteristics *X*, *Y*, and *Z*.
2. Case  $\alpha$  is an *A*.
3. Therefore  $\alpha$  has characteristics *X*, *Y*, and *Z*.

*A* can translate as "revolution," "nationalism," "war," or something else, while *X*, *Y*, and *Z* can constitute necessary conditions, sufficient conditions, standard sequences, correlates, or consequences. A statement in this form can easily reduce to a definition, merely affirming that there exists a set of instances sharing properties *X*, *Y*, and *Z*. The statement need not reduce to a definition, however, since the argument can readily incorporate causal, sequential, or transactional links among the elements. The argument does not assert that all instances of *A* are identical, but it does assert that they share essential properties setting them off from all cases of non-*A*; those essential universals mark any such model as invari-

ant. Analysts often arrive at this sort of argument through empirical comparison of cases  $\alpha$ ,  $\beta$ ,  $\Gamma$ , and so on, searching for the cases' common properties that qualify them all as  $A$ 's. In the domain of large-scale politics, at least, such reasoning so badly describes what actually occurs as to hinder sociological analysis.

"Invariant" does not equal "general." Laws concerning variation sometimes cover a very general range. For an example, consider Boyle's law: at a given temperature the pressure of a certain mass of confined gas varies inversely with its volume. Although we have no well-established sociological laws with the elegance of Boyle's formulation, broad and robust empirical generalizations concerning variation—for example, that over large populations infant mortality declines as literacy rises—abound in different realms of social life. I am not challenging the possibility of more explicitly causal laws of extremely general scope, just so long as they stipulate variation. I am instead challenging the common, if often implicit, claim for essential, invariant universals.

Such claims appear frequently in macrosociology. Take the case of revolutions (for helpful reviews, see Berejikian 1992; Boswell 1989; De-Fronzo 1991; Dunn 1989; Goodwin and Skocpol 1989; Hobsbawm 1986; Keddie 1992; Kimmel 1990; Knight 1992; Outram 1992; Rice 1990; Schutz and Slater 1990; Taylor 1984; Wickham-Crowley 1991; Zimmerman 1983). Generations of scholars have pursued the chimera of an invariant general model of revolution. Fixation on invariant models gives rise to a common but logically peculiar sociological performance we may call "improving the model." It consists of (1) outlining a widely accepted model of phenomenon  $A$ , (2) identifying an instance of  $A$  that fails to fit the model in one or more ways, (3) modifying the model so that it now accommodates the previously exceptional instance as well as those instances that already belonged to its domain. Most often the crucial modification respecifies a condition postulated as necessary in the model's previous version. Thus improving the model expands the claimed scope of the alleged invariance. The procedure is peculiar both because it makes implausible allegations of invariance and because it attenuates whatever empirical grasp the previous model attained. Yet as a reviewer for professional journals I read a half-dozen drafts each year that follow just such reasoning.

Similar reasoning motivates whole books. When Farideh Farhi compares the Iranian and Nicaraguan revolutions that began in 1979, for example, she explicitly sets up the analysis as an extension of Theda Skocpol's *States and Social Revolutions* on the grounds that Skocpol's book is "perhaps the most comprehensive attempt to bring together the new concerns about the role of the state, the structure of peasant communities, and the role of international factors in understanding the processes

and outcomes of revolutions" (Farhi 1990, p. 5). Her methodological declaration runs this way: "The essence of comparative history is to maintain the particularity of each case while accepting that each particularity is shaped by general forces operating at the societal or global level. Accordingly, the intention is to expose these forces as they impinge on quite specific and unique circumstances in the hope of shedding light on historical specificities as well as the changing structures in the larger world-historical context that make contemporary revolutions not utterly unlike 'classic' revolutions but also not totally similar to them" (Farhi 1990, p. 2). Thus all revolutions share attributes *X*, *Y*, and *Z*, even if they differ with respect to a great many other attributes; an effective analysis combines specification of universals with enumeration of particulars.

True to the challenge, Farhi works with a checklist drawn directly from Skocpol: conditions favoring class coalitions against the regime, circumstances promoting the mobilization of those coalitions for revolutionary action, factors making the state vulnerable to attack, and so on. Almost inexorably, this leads her to propose one-for-one substitutes for the factors Skocpol emphasized—for example, Farhi offers the connect-edness and proximity to power of capital cities as a substitute for Skocpol's solidarity of peasant villages. She finally seeks to build a bridge from Skocpol's model to her own by (a) showing how the world development of capitalism has altered class structures since the times of Skocpol's revolutions and (b) attributing more importance to ideology, including religious belief, than Skocpol was ready to concede in 1979.

Such an analysis aims to generalize Skocpol's model rather than to extract from it principles of variation. But it misses the mark: the collapsing agrarian bureaucracy overburdened by international pressures and the autonomous peasant communities aligned against their landlords— keystones of Skocpol's theory—disappear from view, with their replacements in Farhi's analysis by no means members of the same causal categories. Thus Farhi draws useful questions from Skocpol, but in pursuit of those questions tacitly abandons the effort to generalize an invariant model of revolution. Indeed, she has no choice; the model will not, cannot, generalize that far.

Another recent example marks even more precisely the blind alley into which the quest for invariant models has led analysts of revolution. In their excellent compilation on Third World revolutions, Jack Goldstone, Ted Gurr, and Farrokh Moshiri (1991) offer an "analytical framework" that continues the quest for invariance. En route, however, they make two turns that send them in precisely the opposite direction, toward broad and incessant variation.

In his theoretical introduction, Goldstone singles out a trio of recurrent

causes for revolution: "Declining state resources relative to expenses and the resources of adversaries, increasing elite alienation and disunity, and growing popular grievances and autonomy" (Goldstone et al. 1991, p. 49). The list echoes Goldstone's *Revolution and Rebellion in the Early Modern World*, published the same year as the edited volume, but written over many previous years. In that book, Goldstone gives a strong tone of breakdown to his most dramatic statements:

The causes of revolutions and major rebellions operate in ways that seem remarkably similar to the forces that build up to cause earthquakes. That is, in the years before such a revolution or major rebellion, social pressures for change build. Yet the existing social and political structures for some time resist change (even though pressures and deformations may be visible). Suddenly, however, some response to the mounting pressure—a state bankruptcy, a regional rebellion—occurs which weakens that resistance (like a block breaking off along the fault). At that point, there is a sudden release of the pent-up forces and a crumbling of the old social structures—a revolution or major rebellion. More concretely, the Scots and Irish rebellions in Great Britain in 1637–1641, and the state bankruptcy and calling of the Estates General in France in 1789, were themselves responses to the mounting social and fiscal pressures in those societies. Yet these particular events also served to unleash far greater social pressures, which overwhelmed these states and led to revolutions. (Goldstone 1991, p. 35)

Note several features of this statement: its emphasis on sudden collapse in response to long-term change, its claims to generality, its insistence on uniformity rather than variation.

In his first contribution to *Revolutions of the Late Twentieth Century*, Goldstone claims continuity in these terms: "In my work on early modern revolutions, I identify three conditions whose *conjunction* led to state breakdown: fiscal distress, elite alienation and conflict, and a high potential for mobilization of the populace. Although the particular forces that create these conditions may be quite different in contemporary societies than in earlier ones, I believe these conditions remain central to the development of revolutionary crises" (Goldstone 1991, pp. 37–38). Leaping nimbly past the problem of specifying how an observer would know in advance of a revolution's actual occurrence when the three bundles of causes were approaching critical mass, Goldstone immediately concedes that these conditions "may be produced by a variety of forces, depending on how they interact with the institutions and structures in particular societies" (p. 49). Population pressure, that powerful propellant of state breakdown in Goldstone's *Revolution and Rebellion*, now fizzles to a force that "may have either positive or negative impact" (Goldstone et al. 1991, p. 40).

Goldstone et al. also propose three general stages of revolution: state crisis, the struggle for power, efforts at reconstruction. These stages,

however, constitute no verifiable theory; they follow tautologically from the book's definition of revolution as "the forcible overthrow of a government followed by the reconsolidation of authority by new groups, ruling through new political (and sometimes social) institutions" (Goldstone et al. 1991, p. 37). At the end of the first turn, then, our voyagers have little more baggage than the explication of a definition.

Then they arrive at the second turn: a recognition that post-1945 revolutions occur in quite different ways from their predecessors because geopolitical settings, dominant ideologies, and international intervention have changed fundamentally. Indeed, the cases they consider—Vietnam, Nicaragua, Iran, Poland, Afghanistan, the Philippines, Cambodia, Zimbabwe, South Africa, and Palestine—amply demonstrate different patterns from the great revolutions of England, France, Russia, or China. "We think now that a state crisis should not be defined as a specific objective condition but rather as a situation in which significant numbers of elites and popular groups believe that the central authorities are acting in ways that are fundamentally ineffective, immoral, or unjust" (Goldstone et al. 1991, pp. 330–31). Thus one of the three general conditions (relative decline of state resources) collapses into the other two (elite alienation and popular grievances) while injustice, previously invisible, squeezes its way into the argument. By this time, the initial promise of an invariant general model has vanished. Just as the once-hot search for crisp predictors of earthquakes has given way to more general debate about the variable operation of plate tectonics (Acton and Gordon 1994; Girdler and McConnell 1994), the search for unique, invariant properties of revolution has ceded to the conception of a variable field within which revolutions occur.

I have not chosen my example because I think Goldstone et al. are obtuse or empirically mistaken. On the contrary, they have drawn correct conclusions from their evidence: the conditions for revolution are not uniform, but vary from region to region and period to period. The conditions vary as politics in general varies. Because within a given region and period many states share political arrangements, national and international, rough similarities and explicable variations appear in the experiences of connected states with revolution. The search for comparisons close at hand therefore advances understanding, while the attempt to build transhistorical models of revolution is doomed to eternal failure. Goldstone et al. only err in refusing to recognize the general implications for method and theory of their own compelling analyses.

Similar conditions prevail in the study of social movements, nationalism, democratization, and a wide variety of other political phenomena, as well as in the zones of organizational behavior, crime, or urban structure. Over and over sociologists assume coherent, durable, self-propelling



social units, attribute general conditions or processes to those units, invoke or invent invariant models of the relevant conditions or processes, then explain the unit's behavior on the basis of its conformity to that invariant model. It is time to expunge that intellectual procedure.

I am making no plea for historical particularism, much less for epistemological relativism or postmodern linguisticism. I am arguing that regularities in political life are very broad, indeed transhistorical, but do not operate in the form of recurrent structures and processes at a large scale. They consist of recurrent causes which in different circumstances and sequences compound into highly variable but nonetheless explicable effects. Students of revolution have imagined they were dealing with phenomena like ocean tides, whose regularities they could deduce from sufficient knowledge of celestial motion, when they were actually confronting phenomena like great floods, equally coherent occurrences from a causal perspective, but enormously variable in structure, sequence, and consequences as a function of terrain, previous precipitation, built environment, and human response.

For hydrologists, a flood is a wave of water that passes through a basin; a severe flood is one in which a considerable share of the water overflows the basin's perimeter. For our purposes, the equations hydrologists use to compute water flow in floods have three revealing characteristics: they reduce floods to special cases of water flow within basins rather than making them *sui generis*, their results depend heavily on the hydrologist's delineation of the basin, while estimation of the flood's parameters requires extensive empirical knowledge of that basin. Yet the equations embody very general principles, the physics of incompressible fluids in open channels (Bras 1990, pp. 478–82).

Note several implications of the analogy. First, every instance of the phenomenon—flood or revolution—differs from every other one; the test of a good theory is therefore not so much to identify similarities among instances as to account systematically and parsimoniously for their variation. Second, in different combinations, circumstances, and sequences, the same causes that produce floods or revolutions also produce a number of adjacent phenomena: smoothly flowing rivers and stagnant swamps on the one side, coups d'état and guerrilla warfare on the other. Third, time, place, and sequence strongly influence how the relevant processes unfold; in that sense, they have an inescapably historical character. Finally, the events in question are far from self-motivating experiences of self-contained structures; they are local manifestations of fluxes extending far beyond their own perimeters. Floods and revolutions have no natural boundaries; observers draw lines around them for their own analytic convenience. In these regards, they resemble a number of other complex but lawful phenomena: traffic jams, earthquakes, segmented labor mar-

kets, forest fires, and many more. I suppose, indeed, that most interesting social phenomena have exactly these characteristics.

How, then, should we search for the causes of revolutions? Arthur Stinchcombe has long since described one version of the explanatory program: to identify deep causal analogies across detailed features of ostensibly different historical sequences. (The cause of event *X* is the minimum set of antecedents that [1] actually occurred, [2] is generally sufficient to produce events of type *X*, and [3] without which *X* would not have occurred in this setting.)

Concepts are the things that capture aspects of the facts for a theory; they are the lexicon that the grammar of theory turns into general sentences about the world. The argument is that the power and fruitfulness of those sentences is determined by the realism and exactness of the lexicon of concepts, and not by the theoretical grammar. The problem of eliminating false sentences by research, the traditional problem of epistemology, is not as problematic as the problem of having sentences interesting enough to be worth accepting or rejecting. And this is determined by whether or not our concepts capture those aspects of reality that enter into powerful and fruitful causal sentences. (Stinchcombe 1978, p. 115)

For this purpose, Stinchcombe recommends ignoring the "epochal theories" invoked by a Trotsky or a Tocqueville in favor of the causal reasoning by which these thinkers chain together narratives. That means breaking down big events into causally connected sequences of events, and examining each link in the chain. More generally, Stinchcombe advocates a shift of attention away from a priori theorizing toward rigorous examination and reduction of analogies, step by step within causal sequences (Stinchcombe 1978, p. 28). At that level, says Stinchcombe, much of the apparent disagreement between a Trotsky and a Tocqueville dissolves. Great historical analysts employ far more similar causal accounts than their competing epochal pronouncements suggest.

Stinchcombe stresses epistemology, conditions for the generation of knowledge. I am stressing ontology, the nature of that which is to be known. But our programs dovetail. If the social world actually fell into neatly recurrent structures and processes, then epochal theories, invariant models, and the testing of deductive hypotheses would become more parsimonious and effective means of generating knowledge. Because the social world does *not* conform to that prescription, we need other programs on both ontological and epistemological grounds. Our programs converge in the historically embedded search for deep causes operating in variable combinations, circumstances, and sequences with consequently variable outcomes. Most of the work therefore concerns not the identification of similarities over whole structures and processes but the explanation of variability among related structures and processes. In studies of

revolution, the work entails explaining why and how different sorts of social settings produce different varieties of forcible seizures of power over states.

There is hope. Not everyone who analyzes revolutions and related phenomena resorts to invariant models. In a wide-ranging synthesis written before the Soviet Union collapsed, David Laitin sketched a promising theory of variation in the readiness of different national elites to break with Moscow. It argued in part that

the historical dimension that accounts for distinctions between the national movements in the Soviet Union is based upon a single variable—the degree to which elites in the peripheral nationalities received most-favored-lord status in Russia. The historical data show that in the territories west of Moscow, most-favored-lord status was readily granted, even when there were no indigenous lords. Lords in the Turkic areas were often given elite privileges, but they were not given access to positions of high status by right. In intermediate cases like Georgia and Estonia, elite mobility was possible but circumscribed. Certain predictions follow from this: (1) in the most-favored-lord regions there would be powerful symbolic unity among titulars for full independence but a waning of resolve as the conflict of interest among two branches of the titular elites begins to manifest itself; and (2) in the non-most-favored lord regions, the pressure for independence would come more slowly (but once set in motion, there would be unity among the titular elites, with only settled minority populations seeking to slow the process down). (Laitin 1991, p. 157)

Laitin simplifies his work by grouping Soviet regions into two categories, but he clearly invokes a continuous principle of variation. He does not, obviously, provide a complete account of the process by which the Soviet Union collapsed, or by which any particular state emerged from the collapse. On the contrary, he offers a promising candidate for one of the many general principles a properly constructed causal account would invoke.

By his rational choice analysis of conditions for secession, Michael Hechter opens another avenue. Hechter argues that secession results from the intersection of four partly independent processes: (1) creation of regions, (2) mobilization for collective action, (3) development of support for secessionist programs, (4) acceptance of independence by the previously dominant state (Hechter 1992, p. 269). In each case Hechter identifies conditions affecting the extent of two factors: shared or imposed interest in acting to facilitate secession and the capacity to do so. Under the heading of support for secession, for example, he proposes (a) low regional dependence on the host economy and (b) perception of the host state's weakness as major promoters of interest and capacity.

Hechter employs a model of logical concatenation rather than of sequence or political processes; except in the sense of logical necessity, he

neither offers propositions concerning the interaction of his four processes nor postulates a dynamic in which identities, connections, interests, and capacities alter as a function of struggle or accommodation. Within the standard *a priori* limits of rational choice analysis, nevertheless, Hechter's discussion does provide a framework that lends itself to the analysis of choice making in territorial segmentation, civil war, and regionally based revolution (Berejikian 1992; Connor 1987; Gurr 1993; Licklider 1993; Lustick 1993; Strang 1990, 1991). By depicting secession as a highly contingent outcome of interacting political processes, Hechter breaks sharply with invariant models.

In an inquiry that deals more explicitly with structure and sequence than Hechter's does, Peter Bearman (1993) looks closely at changing relations among gentry in Norfolk, England, during the century before the Civil War, which began in 1640. Using the formal techniques of network analysis, Bearman shows that a kinship-based regional structure of power gave way to one based much more heavily on patron-client chains connecting local actors to national centers of power, that gentry experiencing blocked or downward mobility clustered together disproportionately in patron-client networks forming distinctive, antiregime religious identities, and that these shared identities-cum-networks became major bases of political mobilization (Bearman 1993; Bearman and Deane 1992). At no point does Bearman suggest that blocked mobility, patron-client networks, or the other factors he analyzes generally produce revolution; he promulgates no invariant model. But he does provide another illustration of a program that invokes powerful general causes in a particular reconstruction of revolutionary processes.

My own version of that enterprise concentrates on variation within Europe over the last five centuries (Tilly 1993). It distinguishes between revolutionary situations (moments of deep fragmentation in state power) and revolutionary outcomes (rapid, forcible, durable transfers of state power), and it designates as a full-fledged revolution any extensive combination of the two. Chronologies of revolutionary situations in multiple regions of Europe demonstrate the great variation and change in revolutionary processes since 1492. The changes include, for example, an impressive rise in frequency of "national" revolutionary situations: state-fragmenting mobilizations in which at least one party made its claim to state power on the grounds that it represented a coherent, culturally distinct population that was currently receiving unjust treatment.

More important, the revolutionary chronologies illustrate—prove would be too strong a word—how regionally and temporally variable forms of international relations, state power, administrative structure, military activity, extraction, and repression shaped the character of European revolutions, not to mention other forms of political conflict. To the

extent that governmental succession depended on warrior-kings recruited from intermarrying royal patrilineages, for example, revolutionary situations were concentrated at those points when a child or an incompetent came to the throne. In regions of intense commercial activity, for another example, revolutionary situations commonly took the form of urban resistance to princely authority. Revolution turns out to be a coherent phenomenon, but coherent in its variation and in its continuity with nonrevolutionary politics, not in any repetitious uniformity. Its sequences and outcomes turn out to be path, time, and situation dependent, not constant from one revolution to the next.

I do not claim to have been the first to notice this degree of variation; in their practical work, as opposed to their introductions and conclusions, most students of real revolutions proceed as if they were dealing with path-, time-, and situation-dependent phenomena whose individual features—but perhaps not whose totality—can be explained by general political principles, given sufficient information about the context. Nor do I claim that my own recent work provides all the answers to the big questions that students of revolution have been pursuing for centuries. I make only three simple claims: (1) The construction of invariant models of revolution—which remains a major activity among American sociologists—is a waste of time. (2) The poor fit between such models and the actual character of revolutions helps account for the slow accumulation of knowledge on the subject, a problem about which Rod Aya (1990) and James Rule (1989) have recently and properly complained. (3) The same conclusions hold for a wide range of social phenomena, including most or all large-scale political processes.

How, then, can we recognize useful alternatives to invariant models of political processes? Valid analyses of political processes rest first of all on plausible ontologies—representations of what is to be explained in terms of a given process's boundedness, continuity, plasticity, and complexity (e.g., recognizing that nationalism consists of some actors' claim to act authoritatively on behalf of a coherent and solidary people, a claim whose origins, makers, forms, and effects vary enormously over time and space). For variant phenomena, valid analyses specify the fields of variation within which they fall, which means specifying their relation to connected but different phenomena (e.g., delineating the continuum along which lie army-to-army interstate war, covert military intervention, full-scale civil war, guerrilla activities, and terrorism). These valid analyses break complex sequences into events, each of which invokes its own configuration of causes including the cumulative effects of previous events (e.g., separating the conditions under which a cycle of intense social movement activity begins or ends from the conditions under which one movement or another sees some of its demands realized). Their gen-

eral propositions consist of principles of variation for analytically separable aspects of the phenomena under examination (e.g., after noticing that democracy entails broad, relatively equal citizenship that grants citizens substantial collective control over governmental personnel and policies as well as significant protection from arbitrary state action, formulating or invoking separate theories of breadth, equality, control, and protection).

Such analyses immediately yield counterfactuals, specifications of what else could have happened if the causal configuration had occurred differently; thus a valid theory of democratization yields propositions about the conditions for authoritarianism and oligarchy. Within limits, such analyses of variation also yield contingent predictions. By this I do not mean the unconditional predictions of invariant models, in which the appearance of sufficient conditions *X*, *Y*, and *Z* invariably produce outcome *A*, but contingent predictions applying phrases such as "insofar as" to variable conditions, their interactions, and their outcomes. In instances such as Eastern Europe's struggles of 1989, then, we might reasonably hope to specify the fields of variation within which they were occurring, then to anticipate the likely outcomes under various still-contingent conditions.

Mine, then, is no counsel of perfection or cry of despair. For, taken separately, these methodological injunctions have the same comfortable familiarity as invariant models. For all their other weaknesses—vulnerability to spatial autocorrelation, assumptions of boundedness and independence of observations, commitments to linearity, and so on—standard sampling designs and multivariate statistics actually presume some such world. Within these limits, which theories of causality and variation analysts should choose remains just as open as before the elimination of invariant models.

Fortunately, we have no obligation to choose right now; we can wait for results of a productive rivalry, perhaps even of a synthesis from these contentions. For the present, anyone who believes what I have said about invariant general models and who cares about the validity of broad political analyses has plenty of work to do. The crucial theoretical and empirical work should eventually reduce the likelihood that the next major change in world politics will baffle sociologists.

## REFERENCES

- Acton, G. D., and R. G. Gordon. 1994. "Paleomagnetic Tests of Pacific Plate Reconstructions and Implications for Motion between Hotspots." *Science* (March 4), pp. 1246–54.
- Alker, Hayward. 1992. "The Humanistic Moment in International Studies: Reflections on Machiavelli and Las Casas." *International Studies Quarterly* 36:347–72.

- Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, rev. ed. London: Verso.
- Aya, Rod. 1990. *Rethinking Revolutions and Collective Violence: Studies on Concept, Theory, and Method*. Amsterdam: Het Spinhuis.
- Barkey, Karen, and Sunita Parikh. 1991. "Comparative Perspectives on the State." *Annual Review of Sociology* 17:523-49.
- Bearman, Peter S. 1993. *Relations into Rhetorics: Local Elite Social Structure in Norfolk, England, 1540-1640*. New Brunswick, N.J.: Rutgers University Press.
- Bearman, Peter S., and Glenn Deane. 1992. "The Structure of Opportunity: Middle-Class Mobility in England, 1548-1689." *American Journal of Sociology* 98:30-66.
- Berejikian, Jeffrey. 1992. "Revolutionary Collective Action and the Agent-Structure Problem." *American Political Science Review* 86:647-57.
- Boswell, Terry, ed. 1989. *Revolution in the World System*. New York: Greenwood.
- Bras, Rafael L. 1990. *Hydrology: An Introduction to Hydrologic Science*. Reading, Mass.: Addison-Wesley.
- Bratton, Michael. 1989. "Beyond the State: Civil Society and Associational Life in Africa." *World Politics* 41:407-30.
- Caporaso, James A., ed. 1989. *The Elusive State: International and Comparative Perspectives*. Newbury Park, Calif.: Sage.
- Cohen, Jean. 1985. "Strategy or Identity: New Theoretical Paradigms and Contemporary Social Movements." *Social Research* 52:663-716.
- Comaroff, John L., and Paul C. Stern, eds. 1993. "Perspectives on Nationalism and War." Working Paper 163, Center for Studies of Social Change, New School for Social Research.
- Connor, Walker. 1987. "Ethnonationalism." Pp. 196-220 in *Understanding Political Development*, edited by Myron Weiner and Samuel P. Huntington. Boston: Little, Brown.
- Dahl, Robert A. 1989. *Democracy and Its Critics*. New Haven, Conn.: Yale University Press.
- DeFronzo, James. 1991. *Revolutions and Revolutionary Movements*. Boulder, Colo.: Westview.
- Diamond, Larry, and Gary Marks, eds. 1992. "Comparative Perspectives on Democracy: Essays in Honor of Seymour Martin Lipset." Special issue of *American Behavioral Scientist* 35:352-629.
- Diani, Mario. 1992. "The Concept of Social Movement." *Sociological Review* 1992:1-25.
- Di Palma, Giuseppe. 1990. *To Craft Democracies: An Essay on Democratic Transitions*. Berkeley and Los Angeles: University of California Press.
- Dunn, John. 1989. *Modern Revolutions, an Introduction to the Analysis of a Political Phenomenon*, 2d ed. Cambridge: Cambridge University Press.
- Duyvendak, Jan Willem. 1994. *Le poids du politique: Nouveaux mouvements sociaux en France*. Paris: l'Harmattan.
- Eisenstadt, S. N. 1992. "The Breakdown of Communist Regimes and the Vicissitudes of Modernity." *Daedalus* 121:21-42.
- Eyerman, Ron, and Andrew Jamison. 1991. *Social Movements: A Cognitive Approach*. University Park: Pennsylvania State University Press.
- Farhi, Farideh. 1990. *States and Urban-Based Revolutions: Iran and Nicaragua*. Urbana: University of Illinois Press.
- Feshbach, Seymour. 1987. "Individual Aggression, National Attachment, and the Search for Peace: Psychological Perspectives." *Aggressive Behavior* 13:315-25.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, N.Y.: Cornell University Press.
- Girdler, R. W., and D. A. McConnell. 1994. "The 1990 to 1991 Sudan Earthquake

- Sequence and Extent of the East African Rift System." *Science* (April 1), pp. 67-70.
- Giugni, Marco G., and Hanspeter Kriesi. 1990. "Nouveaux mouvements sociaux dans les années '80: Evolution et perspectives." *Annuaire Suisse de Science Politique* 30:79-100.
- Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- Goldstone, Jack A., Ted Robert Gurr, and Farrokh Moshiri, eds. 1991. *Revolutions of the Late Twentieth Century*. Boulder, Colo.: Westview.
- Goodwin, Jeff, and Theda Skocpol. 1989. "Explaining Revolutions in the Contemporary Third World." *Politics and Society* 17:489-509.
- Gurr, Ted Robert. 1988. "War, Revolution, and the Growth of the Coercive State." *Comparative Political Studies* 21:45-65.
- . 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington, D C.: U.S. Institute of Peace Press.
- Gurr, Ted Robert, Keith Jagers, and Will H. Moore. 1990. "The Transformation of the Western State: The Growth of Democracy, Autocracy, and State Power since 1800." *Studies in Comparative International Development* 25:73-108.
- Haas, Ernst. 1986. "What Is Nationalism and Why Should We Study It?" *International Organization* 40:707-44.
- Hall, John A., and G. John Ikenberry. 1989. *The State*. Minneapolis: University of Minnesota Press.
- Hechter, Michael. 1992. "The Dynamics of Secession." *Acta Sociologica* 35:267-83.
- Held, David. 1987. *Models of Democracy*. Stanford, Calif.: Stanford University Press.
- Hobsbawm, E. J. 1986. "Revolution." Pp. 5-46 in *Revolution in History*, edited by Roy Porter and Mikulas Teich. Cambridge: Cambridge University Press.
- . 1990. *Nations and Nationalism since 1789: Programme, Myth, Reality*. Cambridge: Cambridge University Press.
- Kearney, Michael. 1991. "Borders and Boundaries of State and Self at the End of Empire." *Journal of Historical Sociology* 4:52-72.
- Keddie, Nikki R. 1992. "Can Revolutions Be Predicted: Can Their Causes Be Understood?" *Contention* 1:159-82.
- Kimmel, Michael S. 1990. *Revolution: A Sociological Interpretation*. Philadelphia: Temple University Press.
- Kirby, Andrew, and Michael D. Ward. 1991. "Modernity and the Process of State Formation: An Examination of 20th Century Africa." *International Interactions* 17:113-26.
- Knight, Alan. 1992. "Revisionism and Revolution: Mexico Compared to England and France." *Past and Present* 134:159-99.
- Koopmans, Ruud. 1993. "The Dynamics of Protest Waves: West Germany, 1965 to 1989." *American Sociological Review* 58:637-58.
- Krasner, Steven. 1984. "Approaches to the State: Alternative Conceptions and Historical Dynamics." *Comparative Politics* 16:223-46.
- Kriesi, Hanspeter. 1993. *Political Mobilization and Social Change: The Dutch Case in Comparative Perspective*. Aldershot: Avebury.
- Laitin, David D. 1991. "The National Uprisings in the Soviet Union." *World Politics* 44:139-77.
- Lee, Su-Hoon. 1988. *State-Building in the Contemporary Third World*. Boulder, Colo.: Westview.
- Lerner, Adam J, ed. 1991. "Reimagining the Nation." *Millennium* 20:351-525.
- Levy, Jack S. 1989. "The Causes of War: A Review of Theories and Evidence." Pp. 209-333 in *Behavior, Society, and Nuclear War*, vol. 1. Edited by Philip Tetlock et al. New York: Oxford University Press.



- Licklider, Roy, ed. 1993. *Stopping the Killing: How Civil Wars End*. New York: New York University Press.
- Lowi, Theodore J. 1992. "The State in Political Science. How We Became What We Study." *American Political Science Review* 86:1-7.
- Löwy, Michael. 1989. "Internationalisme, Nationalisme et Anti-impérialisme." *Critique Communiste* 87:31-42.
- Lustick, Ian S. 1993. *Unsettled States, Disputed Lands: Britain and Ireland, France and Algeria, Israel and the West Bank-Gaza*. Ithaca, N.Y.: Cornell University Press.
- Mann, Michael, ed. 1990. *The Rise and Decline of the Nation-State*. Oxford: Blackwell.
- Mayer, Margit. 1991. "Social Movement Research and Social Movement Practice: The U.S. Pattern." Pp. 47-120 in *Research on Social Movements: The State of the Art in Western Europe and the USA*, edited by Dieter Rucht. Frankfurt and Boulder, Colo.: Campus/Westview.
- Melucci, Alberto. 1985. "The Symbolic Challenge of Contemporary Movements." *Social Research* 52:789-816.
- . 1989. *Nomads of the Present: Social Movements and Individual Need in Contemporary Society*. Philadelphia: Temple University Press.
- Mitchell, Timothy. 1991. "The Limits of the State: Beyond Statist Approaches and Their Critics." *American Political Science Review* 85:77-96.
- Most, Benjamin A., and Harvey Starr. 1983. "Conceptualizing 'War': Consequences for Theory and Research." *Journal of Conflict Resolution* 27:137-59.
- O'Donnell, Guillermo, and Philippe C. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore, Md.: Johns Hopkins University Press.
- Outram, Dorinda. 1992. "Revolution and Repression." *Comparative Studies in Society and History* 34:58-67.
- Poggi, Gianfranco. 1990. *The State: Its Nature, Development and Prospects*. Stanford, Calif.: Stanford University Press.
- Rice, E. E., ed. 1990. *Revolution and Counter-Revolution*. Oxford: Blackwell.
- Rueschemeyer, Dietrich, Evelyn Huber Stephens, and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.
- Rule, James B. 1988. *Theories of Civil Violence*. Berkeley and Los Angeles: University of California Press.
- Schmitter, Philippe C., and Terry Lynn Karl. 1991. "What Democracy Is . . . and Is Not." *Journal of Democracy* 2:77-88.
- Schutz, Barry M., and Robert O. Slater, eds. 1990. *Revolution and Political Change in the Third World*. Boulder, Colo.: Lynne Rienner.
- Segal, Daniel A. 1988. "Nationalism, Comparatively Speaking." *Journal of Historical Sociology* 1:301-21.
- Sewell, William H., Jr. 1992. "A Theory of Structure. Duality, Agency, and Transformation." *American Journal of Sociology* 98:1-29.
- Somers, Margaret R. 1992. "Narrativity, Narrative Identity, and Social Action: Rethinking English Working-Class Formation." *Social Science History* 16:591-630.
- Starr, Harvey. 1994. "Revolution and War: Rethinking the Linkage between Internal and External Conflict." *Political Research Quarterly* 47:481-507.
- Stinchcombe, Arthur L. 1978. *Theoretical Methods in Social History*. New York: Academic Press.
- Strang, David. 1990. "From Dependency to Sovereignty: An Event History Analysis of Decolonization, 1870-1987." *American Sociological Review* 55:846-60.
- . 1991. "Global Patterns of Decolonization, 1500-1987." *International Studies Quarterly* 35:429-54.
- Tarrow, Sidney. 1993. "Modular Collective Action and the Rise of the Social Movement. Why the French Revolution Was Not Enough." *Politics and Society* 21: 69-90.

- Taylor, Stan. 1984. *Social Science and Revolutions*. New York: St. Martin's.
- Tilly, Charles. 1993. *European Revolutions, 1492–1992*. Oxford: Blackwell.
- White, Harrison. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton, N.J.: Princeton University Press.
- Wickham-Crowley, Timothy. 1991. *Exploring Revolution: Essays on Latin American Insurgency and Revolutionary Theory*. Armonk, N.Y.: M. E. Sharpe.
- Williams, Brackette F. 1989. "A Class Act: Anthropology and the Race to Nation across Ethnic Terrain." *Annual Review of Anthropology* 18:401–44.
- Zimmerman, Ekkart. 1983. *Political Violence, Crises and Revolutions*. Cambridge, Mass.: Schenkman.

# What Can Sociological Theories Predict?

## Comment on Collins, Kuran, and Tilly<sup>1</sup>

Edgar Kiser

*University of Washington*

It may seem as though the articles by Collins, Kuran, and Tilly take very different positions on the possibility of predicting revolutions, ranging from "it cannot be done" to "it already has been done." Randall Collins argues that macrosociological prediction is possible and provides the compelling example of his own successful prediction of the decline of the Russian empire. Timur Kuran suggests that making precise predictions about revolutions is not possible now and may never be possible, due to empirical problems with observing preferences and the complexities of the aggregation of individuals' actions in revolutionary situations. Charles Tilly claims that we have been unable to make successful predictions about revolutions because we have been using the wrong kind of theory.

When we explore their arguments in more detail, it becomes clear that they agree on some issues and that their arguments only partially overlap on others. First, there is agreement among the authors (and among philosophers of science) on the two things that are required for successful prediction: general theoretical models of the relevant causal processes and detailed empirical knowledge of initial conditions. Second, since the authors focus on different levels of analysis, different time frames, and different dependent variables, their arguments are rarely in direct opposition. For example, Collins uses a macrosociological model to make predictions within broad time frames (30–50 years), whereas Kuran focuses on the possibility of using microlevel models and threshold models of aggregation processes to make more precise predictions (within one to five years). Collins's claim that prediction within broad time frames is possible is not incompatible with Kuran's assertion that more precise prediction is not. Furthermore, the arguments about prediction made by Tilly, Collins, and Kuran only partially overlap, because they are not talking about the same explananda. Tilly's arguments apply to all political processes, Kuran focuses only on situations in which preference falsi-

<sup>1</sup> I would like to thank Ron Jepperson and Kevin Neuhauser for helpful comments. Address correspondence to Edgar Kiser, Department of Sociology, DK-40, University of Washington, Seattle, Washington 98195.

fication is present (a subset of political processes), and Collins develops a general theory of state power that has implications for the decline of empires. I will first address issues raised in broad, macrosociological prediction and then turn to the use of microlevel models to make more precise predictions.

#### MACROSOCIOLOGICAL PREDICTIONS

Tilly and Collins seem to come to very different conclusions about making successful macrosociological predictions within fairly broad time frames. Tilly argues that macrosociologists have not been successful in predicting revolutions (or even in explaining them after the fact), because they have tried to apply *invariant* models to different types of revolutions. He claims that, because revolution is a complex and heterogeneous explanandum, it is unlikely that the same set of causal factors will account for all revolutions. In an interesting discussion of the similarities between revolutions and floods, Tilly argues that instead of invariant models of recurrent structures we should attempt to develop general models of causal processes (analogous to the physics of incompressible fluids in open channels) and use them to explain *variations* in outcomes resulting from differences in initial conditions (such as the exact parameters of the river or lake and the amount of water added to it). Sociologists still have very few general models as precise as the physics of incompressible fluids in open channels; we have a lot of theoretical work to do.

Collins's geopolitical model provides an impressive start in this direction. Does his success prove that Tilly's criticisms of current attempts at macrosociological prediction are wrong? Quite the contrary, the success of Collins's prediction supports most of Tilly's prescriptions, although it does suggest that Tilly's criticisms of current macrosociological theory are overstated.<sup>2</sup> Collins does not propose an invariant model of all revolutions but a set of conditionally universal propositions that apply to the rise and decline of empires—and thus to some types of revolutions. Collins's article also illustrates Tilly's points about combining theoretical models of causal processes with detailed data about initial conditions to explain variations in outcomes. In the earlier article in which the prediction was made (Collins 1986), he traces the historical development of the

<sup>2</sup> It is doubtful that the general failure to predict revolutions is due to the use of invariant models. As Kuran points out, it was not only sociological theorists that failed to predict the revolutions of 1989; area specialists and Eastern European political leaders and dissidents were surprised as well. It seems doubtful that most members of the latter groups were using invariant models to analyze the Eastern European political situation. Even within sociology, many scholars of revolution do not use invariant models (some reject all types of general models), and they too were surprised.

Russian empire over several centuries, demonstrating how differences in historical conditions led to different outcomes—in some situations continuing expansion of the empire, in other conditions contraction, fragmentation, and decline.

#### THE MICRO LEVEL AND PRECISE PREDICTIONS

Both Kuran and Collins recognize that the best macrosociological models can offer is prediction within broad time frames; more precise prediction (if it can be done) requires microlevel and aggregation models as well. What types of models are most appropriate for this part of the explanation, and can we expect them to yield precise predictions?

Collins's geopolitical model illustrates a general tendency of structuralist arguments in sociology to leave the microfoundations of their models implicit.<sup>3</sup> Although Collins does not discuss the microfoundations of his macrolevel geopolitical model, he seems to implicitly assume that actors making state policy are rational, self-interested maximizers of some combination of wealth and power. Taking his first proposition as an example, size and resource advantages produce territorial expansion only if policymakers in the larger state recognize and take advantage of the situation. If policymakers instead consistently act on emotions or internalized pacifist values, resource advantages may have little effect on territorial expansion. Making the rational choice microfoundations of the geopolitical model explicit raises some interesting issues. For example, why would rational policymakers systematically overextend themselves beyond the point at which the costs of controlling new territory are greater than the additional resources it produces (principle 5)? The tendency toward overextension is probably a correct empirical generalization, but to explain it theoretically requires more attention to the micro level.

Timur Kuran elaborates on the role of the micro level in revolutions from a rational choice perspective. He suggests that, when regimes are autocratic or issues are particularly sensitive, people engage in preference falsification (expressing public preferences that are contrary to their private preferences), so it is difficult both for people in these regimes and scholars studying them to discover the level and the distribution of discontent. The occurrence of revolutions depends on the nature of peoples' revolutionary thresholds (the point at which the external cost of joining the opposition becomes lower than the psychic cost of preference falsification). The main reason that revolution is so hard to predict is that the effects of structural factors vary depending on the nature and distribution

<sup>3</sup> Collins's model does discuss the microfoundations of action in the "revolutionary days," stressing the role of emotions and symbols.

of existing revolutionary thresholds, and revolutionary thresholds are imperfectly observable due to widespread preference falsification.

Kuran argues that although these factors make it almost impossible to predict revolutions, they do not hinder our ability to explain them after the fact. Although his sharp distinction between prediction and explanation is useful, Kuran's argument understates the difficulty of producing good explanations. Kuran is correct to note that many scholars have developed explanations of revolutions, but some of these explanations are mutually exclusive, and there is often little consensus about which is right. Moreover, it is more difficult to judge the validity of an explanation than the correctness of a prediction, especially if it is an explanation of a small number of cases.

Is Kuran right that it always will be impossible to make more precise predictions about revolutions? I think he is right that precise prediction requires microlevel models and that the detailed empirical knowledge of revolutionary thresholds necessary for predicting revolutions will often be difficult to get. Kuran offers some interesting new measurement techniques as a partial solution to the problem. Let me suggest another potential solution. If we cannot get reliable data on the distribution of preferences (which is the case for most historical work, even in the absence of preference falsification), another alternative would be to make theoretically based assumptions about revolutionary thresholds (Hechter and Kiser 1994).<sup>4</sup> To do this would require models of how structural factors form and transform individuals' interests, including not only their level of support for the regime but their time horizons and their propensity to take risks, as well. With these models, and detailed data on the structural factors affecting individuals (which would be much easier to get than data on their preferences), we may be able to predict revolutions. Since we do not yet have good general models of this sort, I agree with Kuran that *precise* prediction of revolutions is not yet possible.

## CONCLUSION

What are the prospects for future theories to be able to predict revolution? Most of the progress that has been made in our understanding of revolution in recent years has been the result of focusing on two things, states and individuals. First, macrosociologists like Theda Skocpol (1979), Jack Goldstone (1991), and Randall Collins (1986) have explored the political determinants of revolution—including factors like fiscal crisis, losses at war, and military overextension. These arguments have significantly ad-

<sup>4</sup> This strategy would allow us to compare the predictive accuracy of models using different microfoundations and thus yielding different predictions.

vanced our understanding of the causes of revolution and have even produced an important successful prediction. Second, rational choice theorists like Michael Taylor (1988), Siegwart Lindenberg (1989), and Timur Kuran (1991) have dramatically improved our understanding of the microfoundations of revolutionary action and of the mechanisms of aggregation.

The most promising thing about these recent advances is that these theories are for the most part not only compatible but complementary. Most of the recent macrosociological arguments about revolution are materialist versions of structuralism, not at all incompatible with the traditional assumptions of instrumental interests in rational choice. The future thus looks promising—we may be on the verge of a new generation of theories of revolution that will combine macrosociological structural arguments with rational choice microfoundations. These types of composite models could greatly enhance our ability to explain revolutions and perhaps even our ability to predict them.

#### REFERENCES

- Collins, Randall. 1986. *Weberian Social Theory*. Cambridge: Cambridge University Press.
- Goldstone, Jack. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- Hechter, Michael, and Edgar Kiser. 1994. "A Proposal for Conducting Theory-Driven Historical Research." Paper presented at the annual meeting of the American Sociological Association, Los Angeles.
- Kuran, Timur. 1991. "Now out of Never: The Element of Surprise in the East European Revolutions of 1989." *World Politics* 44 (1): 7–48.
- Lindenberg, Siegwart. 1989. "Social Production Functions, Deficits, and Social Revolutions: Prerevolutionary France and Russia." *Rationality and Society* 1 (July): 50–76.
- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- Taylor, Michael. 1988. *Rationality and Revolution*. Cambridge: Cambridge University Press.

# Comment on Kuran and Collins

James S. Coleman  
*University of Chicago*

## KURAN

Stimulated by the East European revolutions of the late 1980s, Timur Kuran has in several papers addressed a puzzling question: Why was everyone, from outside political observers and scholars to citizens in East Europe to the leaders of East European states to dissidents in those states, taken by surprise? His answer lies in the difference between privately held opinions of the citizenry and their publicly expressed opinions and in the conditions under which the gap between the two can remain wide. The essential condition, in brief, is one in which repression of dissent is not only carried out by the regime but is diffused throughout the society, so that people in everyday life profess support for the regime to friends and neighbors, as a means of self-protection. This widespread public expression of support for the regime generates a form of pluralistic ignorance, in which, despite widespread private opposition to the regime, each feels isolated and fearful, afraid that any expression of dissent may be reported to authorities.

In this article, Kuran extends his ideas by introducing the idea that each citizen has a "revolutionary threshold," based on the costs of expressing support for the regime and the costs of expressing opposition to the regime. He points to two principal sources of these costs: The costs of expressing support are in the psychic distress of living a lie, which depends on how strongly opposed one's private opinion is to the regime. On the other side, the costs are the amount of harm one can expect from the regime for expressing opposition.

I find Kuran's thesis an attractive one; it seems to fit well some of the events in some Eastern European societies before and during the 1989 revolution—some, but not all. For example, the gap between private and public opinion is far more descriptive of East Germany in the 1980s than it is of Poland during that same period. In Poland, the microstructure of society had for some time been freed from the fears of expressing opposition to the regime among friends and neighbors. The advent of Solidarity in August 1980, following earlier antigovernment demonstrations in 1968, 1970, and 1976, showed publicly the widespread opposition

© 1995 by The University of Chicago. All rights reserved.  
0002-9602/95/10006-0009\$01.50



to the regime. It is incorrect to argue that in Poland the gap between private and public opinion, generating pluralistic ignorance about the extent of opposition, was an important deterrent to overthrowing the regime.

Yet despite this publicly expressed widespread opposition, the Communist regime was not overthrown until 1989. Why? The answer, I believe, is simple: first, the use of force by the Polish Communist regime under General Jaruzelski, in the form of martial law instituted in December 1980, and second, the fear of a Soviet invasion to restore order and prop up the regime. There was extensive debate among Solidarity leaders in the early days of the movement about how far they should go. The point that always won these disagreements was that they must stop short of whatever would bring about intervention by the Soviet military. The actions of Solidarity did not stop short of inducing reprisals on the part of the Polish government, though they did stop short of inducing Soviet intervention. It remains a matter of debate within Poland whether Jaruzelski was a traitor to Polish freedom by instituting martial law or was the protector of Polish freedom by forestalling a Soviet military reprisal.

Kuran's notion of a revolutionary threshold fits here, but only one of his proposed two determinants of the threshold is important, that is the cost of expressing opposition to the regime. There was little or no gap between private and public opinion and thus little pluralistic ignorance about the extent of opposition.

What, then, brought about the revolt, first evidenced in Poland, and then elsewhere, with the most dramatic events occurring in East Germany? The answer again is simple: removal of the Soviet threat, with Gorbachev's unwillingness to commit Soviet troops to support East European Communist governments. This was most clearly evident when Gorbachev refused Honecker's request to use Soviet troops stationed in East Germany to block the movement of East Germans to the West.

It seems, then, that viewed appropriately, the overturn of repressive regimes in situations like that of Eastern Europe in the late 1980s is far less puzzling and far more predictable than Kuran would have it. That is, it is *contingently predictable*, predictable contingent upon the actions of the repressive regime. In this case, Gorbachev declined to continue to exercise repressive measures to maintain the East European puppet governments. When the threat of repression was removed, the opposition overthrew a toothless regime. This is not to say that the private-public opinion gap did not play a role in determining the revolutionary thresholds in some of the societies, such as East Germany, Czechoslovakia, and Romania. It does suggest, however, that the principal determinant was the presence or absence of the threat of reprisal.

This, of course, does not explain why Gorbachev refused to commit

Soviet forces. This is a different question altogether, with the answers to be sought perhaps in internal problems, economic or military, such as those confronted by the Soviet Union with the Afghanistan fiasco and the economic distress of the 1980s. What this analysis does imply, of course, is that in predicting the overthrow of repressive regimes, the central element to be predicted is the regime's relinquishing the use of repressive force. I suggest that successful prediction of *this* will allow successful prediction of the overthrow of the regime.

#### COLLINS

Note that it is these actions, that is, Gorbachev's refusal to commit Soviet troops in 1989, that constitute a large part of what Collins attempts to predict with his geopolitical theory. He also discusses consequences of these actions. He points to Gorbachev's relaxation on constraints to emigration of Jews to Israel and (in accord with the thesis I have argued above) the response by ethnic minorities within the Soviet Union to this relaxation of repression by spontaneous migrations across republic borders. Collins's discussion is at its most convincing at this point, showing the *consequences* of relaxation of repression, less so in showing the *causes* of this relaxation. I believe it is here, in predicting the causes of relaxation of repression, that theories designed to predict revolutions are weakest. The conditions, military and economic, as discussed by Collins, Goldstone, Skocpol, and others, are clearly important in bringing about state breakdown. However, the decision to relax repressive measures is made by one leader, or at most a few leaders, in repressive regimes. Had Andropov been succeeded by someone other than Gorbachev, the East European governments and the Soviet Union might have remained intact for some time beyond 1989. There is an inherently lower predictability of one person's actions given the conditions confronting that person than there is of the actions of a number of persons, given the conditions confronting them. That is to say more generally that governmental actions are less predictable than the population's actions. If we know the former, then the latter, including an eruption of opposition, are predictable.

#### METHODOLOGICAL INDIVIDUALISM VERSUS HOLISM

There is one major difference between Kuran's approach to prediction of revolutions and that of Collins. Kuran's is an example of methodological individualism, while Collins's is an example of methodological holism. All the variables that Collins uses in his geopolitical model characterize the state or the society as a whole (see, e.g., his fig. 1). There is no attempt to characterize, predict, or explain the actions of those who

initiate a revolt or those who support the revolutionaries, despite the fact that it is their actions that constitute the revolution.

Kuran's theory focuses primarily on the orientations of these persons, the general population, and potential revolutionaries. There are state-level or society-level variables in this theory that are seen to be causes of these orientations (the degree of repressiveness of the regime and the extent to which the regime has successfully instilled fear of friends and neighbors as possible agents of the regime), but the theory moves between the level of the society as a whole and the level of individuals whose actions can constitute a revolt. Perhaps the most ingenious part of the theory is the proposed process through which these orientations of individual citizens can lead to an explosive revolt, a self-propelling process in which the threshold for opposition on the part of one person depends on the number of others already acting in opposition. This produces a positive feedback, which can transform individual "thresholds" into a systemwide conflagration.

It is, I believe, this kind of theory, which moves between the macrosocial level and the microsocial level, rather than a theory that remains at the macrosocial level, that constitutes the more promising avenue for the development of social theory.

# On Grand Surprises and Modest Certainties: Comment on Kuran, Collins, and Tilly<sup>1</sup>

Alejandro Portes  
*Johns Hopkins University*

People who are in the business of making revolutions have long known that there is nothing inevitable in their unfoldings. On the eve of the Russian Revolution, Leon Trotsky reflected with profit on what had happened in Austria in 1848: "When the population of Vienna took *de facto* possession of the city, when the monarchy by this time was on the run, had lost all meaning, when under the people's pressure the last troops were removed from the city, when it seemed that Austrian state power could be had for the asking, no political force was available to take over, . . . the workers were courageous enough to smash the reaction, but not organized enough to become the successors" (Trotsky [1909] 1971, pp. 71–72).

The Bolsheviks learned the lesson well. For a long while, their leaders lashed out at those who would settle comfortably into the arms of historical necessity. When their opportunity came in those fateful days of 1917, they made sure that the power vacuum would not go unoccupied. Political sociologists have long been fascinated by these dramatic events. Unlike historians, who have been generally content with chronicling the conditions and deeds leading to revolutions, sociologists have sought to isolate the deep constants that precede these episodes and account for their emergence. In the process, they have produced persuasive retrospectives. Their record at prospectively anticipating new explosions, however, has been much less encouraging.

The articles that form the core of this symposium take quite different positions with regard to the possibility of prediction of macropolitical events. Kuran argues that we will continue to be surprised by them and offers a rationale based on concealment of private preferences. Collins believes that prediction is possible if you have the right theory and accurate data and offers, as an example, his own prediction of the demise of the Soviet Union. Tilly argues, on the contrary, that the business of

<sup>1</sup> Correspondence may be addressed to Alejandro Portes, Department of Sociology, Johns Hopkins University, Baltimore, Maryland 21240.

prediction of the basis of general invariant principles is a waste of time. In his view, the best that can be hoped for is better understanding of specific events on the basis of variation in the presence, timing, and interaction of deep historical factors.

Before addressing these different positions, I believe it useful to introduce a distinction not discussed explicitly by any of the authors, namely the existence of different types of objects of explanation. At a minimum, three such types can be identified: steady states, trends, and events. The articles in this symposium focus on the last type and, within it, on a particular subcategory, namely large-scale complex events occurring infrequently in different historical periods. This is arguably the most ambitious kind of sociological prediction, but an exclusive focus on it obscures the possibility of similar exercises at less exalted levels. These include institutional and cultural continuities over time (steady states) and gradual transformation over extended periods (trends). Worthy exercises can also include specific events taking place within a restricted sociotemporal space. These latter predictions, which assume a roughly constant cultural and institutional framework, may be labeled "hypotheses of the middle-range," a term popularized by Merton ([1949] 1968, pp. 39–72). It is worthwhile to keep these different types of endeavors in mind while appraising the validity of sociological prospectives on revolution and state breakdown.

Kuran gives us a number of eloquent passages on why such macropolitical events have surprised social scientists in the past and will continue to do so in the future. He bases his otherwise plausible conclusion on a curious rationale. The reason for the continuing surprises is that we do not know the distribution of real private preferences about a political system. If we knew how people under a particular regime really felt, we would be able to predict more accurately the onset of antiregime explosions. Preference falsification not only prevents social science knowledge about revolutions, but retards them since actors under a repressive regime are not aware of their collective real level of discontent. This mechanism, long familiar to social psychologists under the concept "pluralistic ignorance" (Allport 1924; Newcomb 1950, p. 608), is invoked by Kuran as the reason why revolutionary explosions do not occur when widespread disaffection exists.

While preference falsification is an interesting notion that may deserve further attention, it falls short of offering a satisfactory explanation of why revolutions occur or why scholars are surprised by them. Widespread discontent may exist and may be *known* to exist by participants and outside observers alike without producing any significant revolutionary consequences. Half a century before the French Revolution, the divorce between monarchy and nation and the contempt in which the

dissolute Louis XV was held were widely known. In 1751, d'Argenson reported that the government was "neither esteemed nor respected . . . the clergy, the military, the *parlements*, the people, all murmur against and take distance from the government and they are right" (Malet 1922, p. 232). This state of affairs did not prevent the French monarch from dying peacefully in his bed nor, as we shall see, guarantee the triumph of the revolution under his successor.

Similarly, Bolsheviks and other revolutionists in tsarist Russia knew and *counted* on general popular disaffection. For them, widespread opposition to the autocracy was not an empirical fact to be established, but a point of departure. It was enough to recall the language of the St. Petersburg workers' petition to the tsar in 1905: "Despotism and arbitrary rule are strangling us, and we are suffocating. Sire, our strength is at an end! The limit of our patience has been reached; the terrible moment has come for us when it is better to die than to continue suffering intolerable torment" (Trotsky 1971, p. 89).

That petition came on the eve of a giant rally on January 9. There was no "preference falsification" in St. Petersburg but there was no revolution either. Instead, hundreds were massacred as they tried to approach the Winter Palace. So long as tsarist troops were willing to fire, the autocracy was secure. The methodological individualism that underlies this singularly naive position about the importance of aggregate personal preferences is reflected as well in the hypotheses that summarize Kuran's argument. The key one is the second, which predicts that unanticipated regime changes will tend to occur in politically repressive countries. This would scarcely come as a surprise to anyone. Analysts of revolution know well that repressive regimes are most prone to breakdowns and revolutionary aftermaths. That is not the question. Instead, the analytic problem is to identify those historical conjunctures or set of specific circumstances that lead to the demise of these regimes. By itself, aggregated individual disaffection will not lead to that outcome, no matter how apparent it is.

Randall Collins is more optimistic about sociological prediction of macropolitical events. He recounts how, on the basis of geopolitical theory and his own analytic skills, he was able to anticipate the demise of the Soviet Union a full decade before the fall. An impressive achievement indeed, but can it be repeated? Collins tell us that, with a valid theory and good empirical data, accurate predictions of future events can be issued on a regular basis. In his view, the principles of geopolitical theory have invariant validity across cultures and historical periods and, in conjunction with recent theories of revolution such as those of Skocpol and Goldstone, these principles can be used to anticipate future state breakdowns.

Collins's vision opens rosy vistas for the future of macrosociology, but alas there are good reasons to believe that this is only a dream. First, even if valid, geopolitical principles are sufficiently vague to lend themselves to contrary interpretations. Collins himself offers a poignant example. Roughly at the same time that he was making his prediction about the breakdown of the Soviet Union, another analyst, Paul Kennedy (1987), was using the same principles to predict the likely demise of the United States. Collins faults Kennedy for being too bound to his own national circumstances, but this only goes to show that geopolitical analysis varies with the perceptions and standpoint of the analyst.

Take another example. Using the same principles, a third hypothetical forecaster could have predicted that China would be the country to experience state breakdown. China is an "interior state" surrounded by enemies on all sides: the huge Russian empire to the north and west, India and Vietnam to the south, and Japan to the east. It is a relatively poor state for its size and it confronts the same problem that has led to the repeated breakdown of Chinese imperial authority over the centuries: huge borders and diverse and fractious ethnic minorities. The overextension of Chinese central authority would have led to the prediction of imminent collapse once things started to unravel in Tiananmen Square. But history did not happen that way: the People's Army did remove the protesters and shot those offering resistance. Order was restored. Today, the Chinese communist state is regarded as an exemplar of pragmatism and of flexible adaptation to the forces of global capitalism.

The point is not only that geopolitical analysis can lead to different predictions but that the time and place-bound concatenation of forces leading to the events called revolutions are impossible to predict on the basis of general principles. Fifty years before the French Revolution, Minister Bernis confided to a colleague that "only a miracle can get us out of the abyss; our system comes unstuck everywhere" (Malet 1922, p. 232). The trend toward disintegration was evident, but this did not guarantee the outcome. When Robert Jacob Turgot became Louis XVI's minister of finance, a golden opportunity was created to save the ancien regime from itself. Turgot implemented a swift program of reforms aimed at stimulating trade, rectifying the taxation system, and improving state finances. The program was beginning to bear fruit when the minister, one of the most able French economists of the century, was brought down by a conspiracy led by the queen. Had Marie Antoinette not chosen to cast her lot with those affected by the reforms and had the king held firm against her whims, the course of French history might have been altogether different. Even on the eve of the Estates General, Calonne might have saved the day had Louis chosen to back his policies (see Lefebvre 1947, pp. 21–28; Dalberg-Acton 1962, pp. 42–45). Thereafter,

events unfolded in a rapid manner full of unexpected turns, prompting Alfred Cobban's (1965, p. 153) remark: "All through the Revolution we find that theory plays little part in determining policies, though it has played much in their subsequent interpretation."

Collins makes the same point implicitly when he recognizes the imprecision of his own and future geopolitical predictions. An expectation with a 50-year range is not really a prediction about an event, but about a trend. This we can do, whether inspired by geopolitics or some other theory. Steady states and trends are within range of sociological predictions (although they may not always be accurate); sudden macrosocial explosions are not. Awareness of a downward trend may even bring about effective attempts to reverse it. These occurrences that, paraphrasing Merton ([1949] 1968, pp. 475–90), may be dubbed "self-negating prophecies" find a wealth of contemporary illustrations. The elite settlements that rescued the Venezuelan and Bolivian states from a process of state disintegration under internal conflict and runaway inflation come to mind as examples (Burton and Higley 1987; Mayorga 1992). So does the apparent ability of Chinese communist administrators to countermand economic stagnation through selective opening to the West and the capacity of the Cuban revolutionary government under Fidel Castro to forestall—so far—predictions of its demise (Nee 1989, 1991; Walder 1992; Eckstein 1994; Perez-Stable 1993).

These conclusions come close to Tilly's critique of the search of invariant causes of state breakdowns. The identification of such causes as "interelite conflict" or "geopolitical overextension" and the intellectual exercise of looking for their presence in past revolutionary explosions makes for interesting reading and perhaps a more systematic comprehension of historical events. They contribute little, however, to the anticipation of future occurrences for the reasons explained previously. Tilly's critique leaves us, however, with the uneasy feeling that all that is possible to undertake in sociology are case studies and explanatory historical analyses. Despite his protestations to the contrary, this sense lingers because he offers no systematic alternatives to the prediction of huge events. While agreeing with his basic arguments, it seems important to supplement them by considering the other side of the question.

What can sociologists really predict? With reasonable expectation of success, we can predict certain trends at different levels of analysis, from small communities to nation-states. The "constant" causes of revolution identified by historical sociologists may be turned with profit into predictions of trends. For example, efforts to increase taxation of the privileged in an autocratic country can be expected to lead to greater interelite conflict and, hence, greater state vulnerability to popular rebellion (Skocpol 1979, pp. 24–33). The final outcome will remain, however, uncertain.



In *The Fiscal Crisis of the State*, political economist James O'Connor (1973) predicted the financial breakdown of the American state under the dual strains of its accumulation and legitimation functions for sustaining advanced capitalism. This argument (which bears a close resemblance to the geopolitical overextension principle) held true as a trend. However, as did other Marxist analysts of the period, O'Connor confused *trend* with *outcome*. The strains were there and actually increased; the predicted final crisis never materialized.

Since functionalism fell into disrepute, most sociologists have been loathe to focus on steady states and the processes contributing to their stability. Social ecologists, whose training leads them to focus on institutional survival, represent a partial exception but, for the most part, contemporary macrosociologists tend to avoid the issue. Though less exciting than big revolutionary explosions, the continuity over time of normative structures and of institutions such as religion, political parties, and corporate firms offer fertile ground for sociological inquiry. They also permit the formulation of predictions with a reasonable degree of accuracy. Although stable structures are commonly taken for granted, the question of how they arise and what keeps them going provides at least as solid a basis for theoretical development as the analysis of social change.<sup>2</sup>

Finally, events that take place in restricted spatiotemporal contexts are also amenable to prediction. In a number of sociological fields, there has accumulated sufficient knowledge to make possible this kind of bounded predictions. Theoretical advancement in political sociology, the sociology of immigration, and the study of crime and delinquency make it possible to advance fairly accurate predictions about such events as community voting outcomes, the entrepreneurial success of particular immigrant groups, and the presence and growth of criminal gangs. The wisdom of the middle range is evident in these and other fields of inquiry. While predicting the big bang will continue to titillate the imagination and ambition of scholars, it is likely that significant progress toward prediction of events will be achieved within more limited confines.

## REFERENCES . .

Allport, F. H. 1924. *Social Psychology*. New York. Houghton Mifflin.

<sup>2</sup> Recent developments in economic sociology illustrate this point. In a sustained dialogue with institutional economists, sociologists working in this field have examined the origins of firms, the stability of corporate hierarchies, and the development of stable market practices. Significant theoretical development in the "new" economic sociology is due at least as much to the analysis of such states and continuities as to that of processes of change. See Granovetter (1994, 1995), Swedberg (1994), and White (1993).

- Burton, Michael G., and John Higley. 1987. "Elite Settlements." *American Sociological Review* 52 (June): 295-307.
- Cobban, Alfred. 1965. *A History of Modern France*. New York: Braziller.
- Dalberg-Acton, John E. 1962. *Lectures on the French Revolution*. New York: Farrar, Straus, & Cudahy.
- Eckstein, Susan E. 1994. *Back from the Future: Cuba under Castro*. Princeton, N.J.: Princeton University Press.
- Granovetter, Mark. 1994. "Business Groups." Pp. 453-75 in *The Handbook of Economic Sociology*, edited by N. J. Smelser and R. Swedberg. Princeton, N.J.: Princeton University Press.
- Kennedy, Paul. 1987. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House.
- Levebre, Georges. 1947. *The Coming of the French Revolution*, translated by R. R. Palmer. Princeton, N.J.: Princeton University Press.
- Malet, Albert. 1922. *Nouvelle histoire de France*. Paris: Hachette.
- Mayorga, Rene A. 1992. "Gobernabilidad en Bolivia." Pp. 21-128 in *Democracia y Gobernabilidad en America Latina*, edited by R. Mayorga. Caracas: Nueva Sociedad.
- Merton, Robert K. (1949) 1968. *Social Theory and Social Structure*. New York: Free Press.
- Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54 (October): 663-81.
- . 1991. "Social Inequalities in Reforming State Socialism. Between Redistribution and Markets in China." *American Sociological Review* 56 (June): 267-82.
- Newcomb, Theodore M. 1950. *Social Psychology*. New York: Holt, Rinehart, & Winston.
- O'Connor, James. 1973. *The Fiscal Crisis of the State*. New York: St. Martin's Press.
- Perez-Stable, Marifeli. 1993. *The Cuban Revolution: Origins, Course, and Legacy*. New York: Oxford University Press.
- Skocpol, Theda. 1979. *States and Social Revolution: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
- Swedberg, Richard. 1994. "Markets as Social Structures." Pp. 255-82 in *The Handbook of Economic Sociology*, edited by N. J. Smelser and R. Swedberg. Princeton, N.J.: Princeton University Press.
- Trotsky, Leon. (1909) 1971. *1905*. London: Penguin.
- Walder, Andrew. 1992. "Property Rights and Stratification in Socialist Redistributive Economies." *American Sociological Review* 57 (August): 524-39.
- White, Harrison. 1993. "Markets in Production Networks." Pp. 161-75 in *Explorations in Economic Sociology*, edited by R. Swedberg. New York: Russell Sage Foundation.

## Book Reviews

*Critical and Effective Histories: Foucault's Methods and Historical Sociology.* By Mitchell Dean. London: Routledge, 1994. Pp. ix + 237. \$62.50 (cloth); \$16.95 (paper).

Barry Smart  
*University of Auckland*

The field of sociological inquiry is populated by a range of pivotal figures, theories, and methods. One of the central themes of Mitchell Dean's book is that a strong case can be made for adding Michel Foucault to the already considerable list of analysts who have made a distinctive, albeit indirect, contribution to the development of the discipline. Given the already large number of monographs devoted to exposition or criticism of Foucault's work, the prospective reader might be excused for thinking that there really cannot be much left to say. However, as the title suggests, this is not simply another text offering an exegesis of Foucault's work.

*Critical and Effective Histories* does provide a thoughtful reading of the development and significance of Foucault's work, but it is a reading that has as one of its parameters a consideration of the prospects for a reconstituted historical sociology. As the author explains, "Far from making Foucault into a sociologist in disguise, I want here to reflect on *his* historical sense and what it might offer the new historical sense in sociology" (p. 12). The reflections that follow encompass not only the different twists and turns taken by Foucault's work but in addition explore the complex relationships of similarity and difference that exist with the works of a series of major figures.

If, as Dean argues, the impact of Foucault's work on the fields of sociology and history has been oblique rather than direct, it has nevertheless been substantial. Foucault's work exemplifies both the possibility and the value of an alternative intellectual practice, one that Dean terms a "problematizing" practice in contrast to "progressivist" forms of theorizing, "exemplified by the narratives of the Enlightenment, by the nineteenth century positivism of Comte, and by elements of Marx's theory of history and certain interpretations of Weber's conception of rationalisation" (p. 3), and to critical forms of theorizing, associated with the dialecticians of "Western Marxism," the Frankfurt school, and Habermas. Foucault's problematizing intellectual practice disturbs and disrupts the philosophies of history implied in the above. Dean's central question, as he explains in the introduction to the volume (p. 5), is What

Permission to reprint a book review printed in this section may be obtained only from the author.

do Foucault's various "histories of the present have to offer historical-sociological studies?" The answer is clearly, carefully, and convincingly elaborated throughout the text.

The opening chapter briefly addresses the state of sociology and its relationship to history and argues that a new style of engaged historical inquiry, influenced by Foucault's histories of the present, has emerged "in contradistinction to . . . a social theory and sociology that seeks to sacrifice historical intelligibility in favour of models, types, and method . . . [and] a conventional historiography that seeks to abandon concepts in the quest for exhaustive reconstruction of the past" (p. 21). Dean elaborates the argument for critical historical-sociological studies by distinguishing between Foucault's interest in developing histories of the present and related conceptions to be found in the respective works of Fernand Braudel and the *Annales* school and Norbert Elias. The pursuit of what Habermas has described as the problem of "presentism" in Foucault's work leads Dean on to a discussion of Foucault's reading of and relation to Kant. As Dean remarks, the "problem of the present may be that which Foucault discovers in Kant. . . . But this problem is radically incomplete unless the further questions are asked: Which today and yesterday, and for whom?" (p. 52).

Reflections on Kant's philosophy and the idea of enlightenment provide an appropriate foundation for three chapters addressing Weber, rationality, and the subject, the "rationalisation thesis" attributed to Weber, the works of Horkheimer and Adorno, and Habermas, respectively. Dean draws attention to the philosophical anthropology intrinsic to Weber's work and argues that, notwithstanding "points of contact and overlap" with the work of Foucault, "at a methodological and meta-historical level their positions are not simply different but fundamentally incommensurate" (p. 60). Likewise, Foucauldian and Frankfurt critical analyses are distinguished "by their terrain, their form and conception of critique and hence of analysis, . . . their conception of the subject" (p. 114), and their conceptions of reason.

Following a carefully measured rebuttal of the criticisms of Foucault's work presented by Habermas, Dean moves away from methodological concerns and turns to a consideration of Foucault's work on relations of power, governmentality, ethics, and practices of the self. Contrasting Foucault's work with historical-sociological writings on the state, Dean argues that "Foucault has opened up the analysis of the state in a way state-focused theory and analysis could not" (p. 153).

Dean's fine discussion of the value of Foucault's work for critical historical sociological analyses of forms of rationality, relations of power, and techniques of the self will undoubtedly prove to be immensely valuable not only to those interested in the work of Foucault but also to analysts concerned to promote the present relevance of sociology.

*Durkheim's Philosophy of Science and the Sociology of Knowledge: Creating an Intellectual Niche.* By Warren Schmaus. Chicago: University of Chicago Press, 1994. Pp. x+314. \$50.00 (cloth); \$24.95 (paper).

Mark A. Schneider  
*Southern Illinois University at Carbondale*

Warren Schmaus's book is aimed at two distinct sets of readers, and it should prove intriguing to a third. It directly targets Durkheim scholars and students of the controversies swirling around science studies. Apt to benefit collaterally are casual students of Durkheim and/or sociological theorists. Let me begin with this third audience.

Casual students like me may assume that a bad idea like the collective consciousness—a veritable swine among pearls—must have been lightly held by Durkheim and utilized largely for its strategic value. Schmaus demonstrates the contrary. Durkheim's philosophy of science, Schmaus shows, made collective representations seem attractive, indeed almost inescapable. And the trouble Durkheim experienced in trying to explicate and explore them largely determined the steps in his intellectual life. For instance, Schmaus argues that by the second edition of *Les règles* in 1901, Durkheim had concluded that collective representations in advanced societies behaved in ways too complex to be tractable. He turned to simple societies in hopes of greater understanding, yet warrant for a distinctively social level of consciousness eluded him even there.

Followers of controversies in science studies will find in Schmaus's book an aggressively internalist explanation of Durkheim's intellectual trajectory, as well as an intriguing examination of his boundary work, which aimed at establishing a place for sociology and for Durkheim himself among his colleagues on *L'Année Sociologique*—the “intellectual niche” of the book's subtitle. Schmaus argues that we can better explain Durkheim's development with knowledge of his realist philosophy, his boundary concerns, and the concrete problems he faced in each ensuing empirical enterprise than we can by appealing to “external” factors. Schmaus specifically faults explanations based on scientists' self-interests, explanations that are unable to account for the cool reception some of Durkheim's notions received within the circle of *L'Année Sociologique*.

Durkheim scholars will find Schmaus stepping quite hard on a wide assortment of toes. Schmaus comes at Durkheim from a fresh angle—not the history of sociology but the philosophy of science and its history—and illuminates Durkheim's development. His first two chapters outline Durkheim's stance as a realist and essentialist who adopted and then further developed “Mill's methods of elimination and Bacon's notion of a crucial test” (p. 7) and used them to create his intellectual niche. The remainder of the book is devoted to detailed explications of Durkheim's

major works, showing how they evidence and affect his philosophy of science. Schmaus argues that Durkheim's empirical work developed seamlessly over the course of his career; he thus challenges those who see "breaks" or fundamental changes in Durkheim's position. Further, he argues that basic misunderstandings of Durkheim's philosophy of science have caused interpreters to misread him badly, especially as regards the causal argument in *The Division of Labor*. Obviously, the book is must reading for Durkheimians.

Let me note two problems, both incidental in the larger scheme of things. Schmaus concludes that *The Elementary Forms* lacks an explanation for "the sacred character of the soul and the totemic principle" and that this "is a major weakness in [Durkheim's] theory" (p. 224). Yet Durkheim describes the "collective action . . . [that] arouses the sensation of sacredness" and attaches it to the totem in section 3 of chapter 7. This is the well-known "corroboli" passage that links the distinction between sacred and profane to the difference between an effervescent "carnival" time and the mundane remainder of the year. Admittedly, the passage is somewhat out of place in the organization of the book, but it certainly leapt out at me upon first reading. Schmaus's inattention to it seems puzzling, unless we attribute it to his provenance as a philosopher of science. Similarly, while a sociological explication of *Suicide* would probably mention Durkheim's notion of a "coefficient of preservation," the emphases of Schmaus's argument cause him to ignore it. Thus, this book is not (nor was it designed to be) the place to send a student for an overview of Durkheim's major works.

A second minor difficulty lies in Schmaus's suggestion that because Durkheim provides a way for "meanings to play a role in scientific explanation," his thought offers "a model for resolving the current crisis in sociology in which theoretical and empirical studies are going their separate ways and even empirical investigations are divided between interpretive and more positivist approaches" (p. 4). For reasons too complex to detail, I think that this places a quite improbable burden on Durkheim's thought and that it registers a certain tone deafness to recent debates, though we may excuse this in view of Schmaus's primary allegiance. In any event, the benefits we derive from his point of view vastly outweigh the minor problems it may occasion. I found his book highly stimulating.

*Language and Reason: A Study of Habermas's Pragmatics.* By Maeve Cooke. Cambridge, Mass.: MIT Press, 1994. Pp. xii + 207. \$30.00.

James J. Chriss  
*University of Pennsylvania*

This book is one of a series from the MIT Press dedicated to studies of contemporary German social thought. Jürgen Habermas is without a

doubt one of the most important and prolific social thinkers in contemporary philosophy and social science—German or otherwise—so it is understandable that Habermas's work and analyses of his works dominate this series.

Maeve Cooke's aim in *Language and Reason* is to provide a general introduction to Habermas's program of formal pragmatics. What Habermas means by formal pragmatics is his own effort at reconstructing the universal principles of understanding that undergird, or at least make possible, everyday communicative practices. To get into this admittedly difficult material, Cooke spends the first chapter summarizing Habermas's conception of communicative action. She draws from a wide range of Habermas's writings, especially his mammoth *Theory of Communicative Action*. Although Cooke's interpretation is solid, here lies one of the drawbacks of such exegeses of Habermas: because Habermas's arguments are so multifaceted, so overarching, so difficult, any secondary analysis must, for the benefit of uninitiated readers, spend time attempting to simplify, or at least highlight, the most relevant parts of Habermas's corpus. Most secondary authors are competent to do this, but after all is said and done it probably makes more sense simply to go to Habermas to learn what is meant by "formal pragmatics," "communicative rationality," "validity claims," "dialogical-fallibilism," and so forth.

This is not to say that Cooke's book should be summarily dismissed. There are a few genuinely insightful contributions to our understanding of Habermas's pragmatics and his conception of communicative rationality contained herein. In chapter 2 Cooke is correct in noting that Habermas has distanced himself from his original notion of the "ideal speech situation" and now speaks more carefully of the strong idealizations of "discourses." The latter refers to forms of argumentation that not only satisfy the idealizations of everyday talk (e.g., assumptions that everyone engaged in communicative exchanges agrees on the meaning of terms) but also that point to a rationally motivated consensus on the *universality* of the validity claims of propositional truths ("constatives"), normative rightness ("regulatives"), and subjective truthfulness ("expressives"). Further, this consensus on the universality of the three validity claims attached to discourse makes sense only within the context of modern or "postconventional" societies, that is, societies that rely not so much on tradition or other rigid conventions for deriving meaning or understanding, but on the principle that no argument is regarded as irrelevant or exempt from challenge.

The fact that participants in discourse raise three validity claims simultaneously and must be prepared to defend their claims via rational argument means that the postconventional subject has become "decentered." Persons can no longer legitimately draw from only one sphere of experience or authority (e.g., science, aesthetics, practical knowledge, religion) when making claims about the rightness, truthfulness, or truth about the

world (chap. 3). As Cooke points out, this capacity of actors to "move between worlds" is for Habermas "the formal-pragmatic underpinning of the idea of the formal unity of reason" (p. 92).

Cooke provides in chapter 4 a discussion which clarifies the terminological distinction between *Verstandigung*—the process of reaching agreement—and *Einverständnis*—the state of having reached agreement. Habermas's notion of the strong idealizations of discourse might give critics the impression that he disregards the disharmonies and disagreements that routinely punctuate everyday talk, interested as he is in the idea that consensus (*Einverständnis*) is anchored in the structures of communication (p. 112). Cooke's clarification of Habermas's point is useful insofar as it emphasizes actors' orientation toward rationally motivated agreement as to what is said (*Verstandigung*) rather than viewing actors as somehow magically imbued with the ability to reach agreement.

In the end, Cooke suggests that Habermas's linguistic account of solidarity in postconventional society—namely, the bonds that hold members of a democratically self-regulated lifeworld together via this rationally motivated agreement to reach understanding—is on its own unconvincing, but warns that abandoning the project gives at least tacit approval to the nihilist position which stresses that there is no basis for deriving normative standards in a postmodern world where universally shared notions of the good life no longer seem possible (p. 166). I wholeheartedly agree with Cooke on both counts.

*The Sociology of Deviance: An Obituary.* By Colin Sumner. New York: Continuum, 1994. Pp. xii + 340.

Sudhir Alladi Venkatesh  
*University of Chicago*

*The Sociology of Deviance* is an ambitious essay. It is ambitious in its attempt to write a history of the development of modernity in the West through a sociology of knowledge. It is an essay, as opposed to a rigorous empirical or theoretical study, because it is far more synthetic and exploratory than analytic. These are perhaps the two greatest contributions the book makes to a loosely coherent sociological trajectory, for the research on "deviance" is as varied and complex as the "modernity" which it evidences and has not been subjected to a serious overview. Colin Sumner's objective in this work is to "explore the theoretical matrix that held the sociology of deviance together and gave it a coherence" (p. viii). As he states outright, he is less concerned with performing a microscopic examination of empirical studies than he is with constructing a "history of the social logic of ideas" that motivated and guided research on deviance from Durkheim to Stuart Hall and Michel Foucault.

Sumner contextualizes his arguments within the dominant culture, politics, and social thought of the three historical periods that produced



distinctive approaches to societal deviations from "normality." His argument begins with the social upheavals at the close of the 19th century and with the work of Émile Durkheim. Turn-of-the-century industrialization, urbanization, and entrenchment of mass production disrupted the traditional bonds of society and, more importantly, gave rise to novel social practices that were not understandable or controllable through extant forms of legal and moral censure. "Everything seem[ed] pregnant with its contrary" (p. 3), wealth and normality could be found alongside urban poverty and "doubt." It was the sociological theory of Durkheim, Sumner argues, that best explicated the implications of such contradictions for social order and the development of individual consciousness. It was Durkheim who both analyzed the causes of anomie and warned his peers that the proper response by a healthy society to individual "deviation" was not outright repression, but instead was a combination of tolerance, censure, and legally driven deterrence and sanction. Durkheim's writing on sociology, which influenced the development of the Chicago school in the twenties and thirties, and the parallel developments in psychiatry would come together to forge the "sociology of social deviation," the first major theoretical nexus guiding deviance research.

Sumner then takes his methodology into the postwar era to outline the transformation of deviance research from "social deviation" to "social control" and "labeling" frameworks (e.g., Goffman, Becker, Matza). Again, he weaves together large-scale historical movements (e.g., the post-war economic "boom") with aesthetic trends (e.g., abstract expressionism) to demonstrate that the changing academic approaches to social deviance can be understood only in the context of broader, more profound developments in societal organization and ideological (self-)representation.

Sumner concludes by analyzing the period 1968–75, "the final collapse" of deviance research. At this time, "the concept of deviance would come home to Europe" (p. 262), with (primarily) British social scientists rescuing deviance from the functionalism, positivism, and rampant empiricism of their American colleagues and reinvesting it with political chutzpa and theoretical sophistication. In doing so, Sumner argues, they effectively pronounced the end of "deviance" research: deviants were not biologically inferior, psychologically imbalanced, or the products of a pathological subculture; instead, they were voicing their (politicized) displeasure with dominant society. Thus, orthodox (depoliticized) criminological techniques were inapplicable.

One can take issue with many of Sumner's claims: his positioning of Goffman and Becker as foundational theorists for other "labeling-oriented" researchers is somewhat tenuous; his treatment of artistic movements is sketchy and is not well justified. However, the most important point of critical interrogation lies in Sumner's objective of building bridges between a disciplinary trajectory and the larger social milieu within which it arose, matured, and died. To make these connections, Sumner crosses a wide terrain, at times clumsily and with no apparent theoretical guidance. Though he supposes himself to be a "neo-Marxist

or Marxisant" (p. 303), the reader is hard pressed to locate his adherence to a dialectical or historical materialist framework for the study of intellectual history. A well-developed justification for his metatheoretical categories would have been welcome.

Despite these criticisms, Sumner's book remains an excellent essay that should be required reading for scholars engaged in deviance research. The sociology of deviance in the United States tends to be less historical and less articulated with many of the sociological traditions discussed in this book. Thus the author provides an excellent example of the insights that can be gained by widening one's theoretical and methodological horizon.

*Illegitimacy, Sex and Society: Northeast Scotland, 1750-1900.* By Andrew Blaikie. Oxford: Clarendon Press, 1994. Pp. xiv + 268. \$55.00.

Peter Becker

*German Historical Institute*

"Perhaps illegitimacy was neither a positive revolt, nor a negative, fatalistic reaction: it was more of a compromise, a fact of life" (p. 231). With this remark about the nature of illegitimacy, Andrew Blaikie concludes his elaborate study of premarital sex in a rural area of Scotland. Though the title of the book suggests broader coverage, Blaikie presents a case study of one rural parish (Rothiemay) whose demographic and social characteristics he sets in the context of all of Northeast Scotland.

Contextualization is certainly not missing in this book. Blaikie studies illegitimate childbirth with constant reference to the empirical evidence in other parts of Great Britain and Europe. He is well informed about historical demographic studies, at least in English. He did not write, however, only another study on premarital sexuality. His work is unique in the way it combines several levels of argumentation to explain and analyze the upsurge of illegitimate childbirth that was primarily a European phenomenon and did not spare the parish of Rothiemay.

Only when statistics became available in 1858 did the bourgeois discourse about sexual misconduct shift its focus from the urban working class to farm servants. Blaikie argues that philanthropic attempts to improve the moral standards of servants had to fail because there were no adequate means for social reform projects (p. 34). He convincingly shows that reform projects were often based on an idealistic image of "traditional" families (p. 62). Blaikie derives from this analysis the main categories of his argument: housing shortages, accommodation of single servants, breakdown of parental control due to high mobility, cultural differences between middle-class observers and farmers and/or servants (pp. 40-41). These social and cultural conditions are the focus of both the contemporary philanthropic discourse and Blaikie's own narrative.

Before proving the validity of his argument, he attempts to apply

the two most influential models for interpreting changing illegitimacy rates—the courtship intensity model and the concept of a bastardy-prone subsociety (P. Laslett, *Bastardy and Its Comparative History* [London 1980])—to his own skillfully assembled evidence. Neither model was, however, found to be adequate for the Scottish experience. Rothiemay's illegitimate childbirth rate was demographically not connected closely enough to courtship to confirm the courtship thesis (see esp. pp. 93, 115). With a close reading of the reproductive careers of mothers of illegitimate children, Blaikie also questions the applicability of the bastardy-prone subsociety. Neither repeaterdom (bearing children out of wedlock repeatedly) nor the spread of illegitimacy within kin networks was found to be significant enough to substantiate the model (see p. 108).

With painstakingly reconstructed biographies of unmarried mothers (taken from census enumeration) and from the records of the Poor Law administration, Blaikie again returns to his own narrative (p. 122): With reference to the social context of bearing children, he takes up issues like the responsibilities for raising illegitimate children that were already discussed by M. Mitterauer in several publications (esp. *Ledige Mütter* [Munich: Beck, 1983]). Blaikie describes the reaction of the kin group as crucial, since bastard bearing should have been pertinent at a point in the life cycle when parental control is absent and young men and women worked and lived away from home (pp. 154–55). To maintain employment during their most productive years, it was necessary for unmarried mothers to have someone take care of the child. With an analysis of census records, Blaikie suggests that there was a link between the availability of parents to raise the illegitimate children and high rates of illegitimacy (pp. 141 ff.). He even considers the positive impact of illegitimate childbirth on the family of the unmarried mother (p. 222) and on the reproduction of the agrarian workforce (p. 231)—an argument that can also be found in studies dealing with German communities (D. Sabean, “*Unehelichkeit*,” in *Klassen und Kultur*, edited by R. Berdahl et al. [Frankfurt am Main: Syndikat, 1982]).

Blaikie does not give in to the temptation to downplay the hardship for women associated with illegitimate offspring. Even though women did receive support from their kin networks, in several instances they were deserted by the father of their child. Blaikie argues that their situation was made easier, however, by the relatively sympathetic view that the welfare institutions in Rothiemay had toward them (p. 183). Nevertheless, “most unmarried women remained trapped within the constraints of poverty which faced many of their class” (p. 227). Their situation was all the more problematic since there existed a highly gendered approach to disciplining sex offenses within the communities. Blaikie proves that fathers of illegitimate children increasingly evaded responsibility and escaped punishment (see pp. 185, 191). With the records of the kirk sessions, Blaikie estimates the degree of acceptance of premarital sex within the community. Since the families of the elders of the kirk session, who were the guardians of morality in the local community,

were infected by the poison of illegitimacy, he concludes that illegitimacy was accepted (p. 205)—at least as a “fact of life.”

Blaikie is very successful in relating the three different sources of analysis—contemporary discourse, demographic evidence, and individual biographies—to each other. He presents an interesting and nicely written account of illegitimacy, but not of sexuality. He uses scattered clues about sexuality, but owing to the silence of the sources he fails to take the difference between sexuality and reproduction into account more systematically. In his narrative he presents many factors that were seen as being responsible for the existence and high degree of illegitimacy. He refuses, however, to integrate them into a larger narrative, thus avoiding what he calls the “tyranny of grand narratives” (p. 231). In this respect, he gives us no new account about the hidden drives behind the rising illegitimacy rates in the 19th century. He provides us, though, with a complex picture that describes the social fabric of rural Scotland probably better than any all-encompassing narrative.

*Gay New York: Gender, Urban Culture, and the Making of the Gay Male World, 1890–1940.* By George Chauncey. New York: Basic Books, 1994. Pp. xii + 478. \$25.00.

Vern L. Bullough  
*California State University, Northridge*

It is George Chauncey's thesis in *Gay New York* that gay life was more integrated into the life of New York City in the first part of the 20th century than it was in the period immediately following World War II. There were, however, periods of greater and lesser public tolerance of gayness, and a secondary thesis of this book is that one aspect of gay life, the bars, declined in importance with the end of prohibition. This was in part due to the election of Mayor Fiorello La Guardia and his reform coalition, but it was mainly due to the establishment of the New York State Liquor Authority (SLA), which was given control over bars. Bar owners could lose their license if their bar seemed to be a place of solicitation or even if gays were too obvious. Bars were inspected often and those whose clientele was judged to be homosexual were raided. This led to a class split in the gay community, since the SLA favored bars whose customers seemed straight over bars whose customers were campy. Since, however, most bar owners could not pay fines or fight a threatened loss of license, organized crime was able to gain a dominant position in the gay bar business.

The center of the gay world moved from the Bowery (where it was based in 1900) to Greenwich Village and Harlem in the 1920s. The Village was the place of residence of many gays and the businesses catering to them. This was, Chauncey implies, what made the Village different from the rest of the city. Harlem, on the other hand, was primarily a

center for gay social life, although gays who were black also lived there. It was only in the Harlem social centers that the races (whether gay or straight) could intermix; ordinarily gay life was otherwise racially segregated.

Baths were also centers of gay life, since several of them catered almost exclusively to male homosexuals. These continued to operate even when gay bars came under attack.

The major social event of the gay season was the costume ball, which was sponsored by a black lodge of the Odd Fellows and held in Harlem. Generally the black community in New York (as in San Francisco and Chicago) proved more tolerant both of its own gay members and of gay outsiders than did the white community. There might well have been hostility beneath the surface, and Chauncey reports the efforts of a major black minister to rid the community of its homosexuals. Chauncey also hints at the influence of the white gay community in popularizing the Harlem Renaissance and describes the contacts among the artists and writers associated with the renaissance and the gay community, but this is not a major theme of the book.

Instead he concentrates on lifestyle, examining the popular press and the black press, various vice commission reports, and the detailed diaries and reports of members of the gay community. He also relies upon interviews published by George Henry (*Sex Variants: A Study of Homosexuality* [New York: Hoeber, 1941]). New York City was a center of American gay life and in Chauncey it has finally found a recorder. New York was not, however, the only center, and I hope that this study will motivate others to examine gay communities in Chicago, San Francisco, and New Orleans, where the sources are equally rich.

In sum, Chauncey gives a strong dissent to those who believe that the gay movement began with the Stonewall riots in 1969. There is a tremendous amount of information available, and Chauncey's studies mark a new beginning in ferreting some of this out. The major failing of the work, which Chauncey recognizes, is that it says little about lesbians. One reason for this is that, like most women, they simply did not get the attention men did. But this study demonstrates that information is available, and I hope it will be further explored.

*Orpheus and Power: The "Movimento Negro" of Rio de Janeiro and São Paulo, Brazil, 1945-1988.* By Michael George Hanchard. Princeton, N.J.: Princeton University Press, 1994. Pp. x+203. \$29.95.

Richard Graham  
*University of Texas at Austin*

Power relations between races in Brazil have long attracted the attention of foreign observers, especially North American ones, even when these writers did not acknowledge that power was what was at issue. In the

19th century scores of travelers reported on the supposedly amicable relations between slaves and masters and between free blacks and whites. Brazilian elites were quick to agree that, indeed, theirs was the land of harmony among the races. To be sure, beginning in the 1890s Brazilian intellectuals, influenced by a "scientific" racism and the works of Darwin, Le Bon, and Gobineau, expressed shame and regret in the face of their nation's mixed-race heritage and large proportion of African-Brazilians. But in the 1930s they and other white leaders returned to the theme of Brazilian exceptionalism, touting their country as the land of racial democracy. This view of Brazil has been thoroughly debunked in scholarly opinion over the last 30 years or so, but it remains deeply entrenched in elite and popular opinion even among blacks.

The evidence of discrimination lies all about: black and mulatto life expectancy, education, housing, and income are vastly inferior to that of whites. Yet these differentials are typically written off to class structure, to the poverty of so many, whether white or black, to the inability or unwillingness of the Brazilian dominant classes to redistribute wealth, and to general economic underdevelopment—but not to racism. That residential segregation is lacking or at least limited, that some blacks and mulattos have made it into the middle and even, a very few, into the upper classes, that so many poor whites live and work alongside blacks, that this is a strongly hierarchical society even for whites, where the ideology of egalitarianism is a recent import and political authoritarianism has been almost the norm—all these facts are adduced to explain away the disparities between whites and nonwhites. It is taken for granted and commonsensical that, racially, equality reigns.

A few African-Brazilians, however, have perceived the racial dimension of their situation and raised their voices in protest. They have formed organizations, staged rallies, organized demonstrations, and published newsletters, but so far they have only begun to be heard. In a country with the largest black population of any nation in the world except Nigeria, movements of black solidarity have been few and far between, short-lived, and until now ineffective.

Why does the myth of racial democracy persist? How is racial domination politically constructed and maintained? Where are its limits? How is it challenged and contested? These are the questions raised by Michael Hanchard's path-breaking study. No one, whether in Portuguese or English, has previously undertaken the task of unraveling the sinews of that firm control, or of examining in a systematic fashion the dynamics of black mobilization.

In carrying out this task, Hanchard appeals to the notion of cultural hegemony as outlined by Antonio Gramsci but applies it to racial rather than class domination. The values and paradigms that help maintain the power of the dominant group are disseminated throughout society by schools, media, religion, and the law to such an extent that they are accepted by the vast majority, even by those who are dominated. Hanchard takes sharp issue with critics of Gramsci, especially with James C.

Scott (*Weapons of the Weak* [New Haven, Conn.: Yale University Press, 1985]), alleging that they have distorted Gramsci's thought in order to attack it. Although Hanchard perhaps does some of the same thing to Scott, he succeeds in drawing our renewed attention to Gramsci's notion of how such a hegemony may be broken and how a counterhegemony is constructed, while at the same time noting that many of the subaltern do not necessarily have a private agenda that contradicts that of the dominant actors.

Hanchard interviewed some 60 black activists in Brazil to explore their attitudes on racial politics and how, based on incidents of their personal lives, they arrived at these positions. By combining "anthropological" (p. 27) with political science techniques, he sought to trace their formative experiences and the way in which they were able to break through the restraints imposed by hegemonic culture. He also examined newspapers, personal archives, and the scholarly literature to delineate the techniques of cultural dominance. After summarizing the history of that dominance, he traces the course of attempts to raise black consciousness in the past, culminating in the theater and countertheater that marked the centenary of abolition in 1988. The result is a book rich in insight and full of striking detail.

Hanchard's point of view is that of a North American sympathetic to the goals of his interviewees, all of whom are black activists. The result is a criticism, both overt and implied, of the nonactivist black community and even of the black consciousness movement itself for its failures to create a coordinated mass-based organization, presumably like those in the U.S. civil rights struggle. This is a committed book, not a distanced analysis. Hanchard believes that, for the African-Brazilians, black identity *should* be more important than class or gender or religion, and he decries their failure to see this truth and join a disciplined and coherent movement. By implication, at least, he has little sympathy for those blacks who believe altering the class structure or improving the lot of particular workers through labor organizing is a more immediate goal than exposing and overcoming racism. He bemoans the internecine struggles among black leaders and their failure to present a united front, as well as their infatuation with African-inspired cultural expressions and their preference for symbolic protest rather than concrete strategies for political change. But Brazilian blacks will follow their own course, shaped by the particularities of their own history, and their achievements need not be measured by a North American model. Nor should we seek to impose our program on them. Since the author's stand is clear, however, there will be no confusion here: Advocacy enlivens the book and makes it all the more important both to the specialist and to the general or classroom reader.

*Urban Unrest in the Middle East: A Comparative Study of Informal Networks in Egypt, Iran, and Lebanon.* By Guilain Denoeux. Albany: State University of New York Press, 1993. Pp. x+310. \$17.95.

Nader Sohrabi  
*University of Chicago*

Concentrating on Egypt, Iran, and Lebanon, Guilain Denoeux examines the effects of rapid urbanization on the unrest that has beset some of the most troubled countries in the Middle East. Central to Denoeux's analysis of urban unrest is the role played by four types of informal networks: patron-client, occupational, religious, and residential. Starting with a useful historical discussion, Denoeux points to the long legacy and continued relevance of informal networks to the political and social life of Middle Eastern cities. The main concern of the book, however, is with the present, and it is here that a comparison with the first half of the 20th century points to an important anomaly: Egyptian and Iranian politics during the 1940s and 1950s appeared much more modern than the politics of recent decades.

Denoeux resolves this paradox by looking at the rise and fall of, and the competition between, formal and informal networks. Prior to the 1950s, newly emerged formal networks such as political parties, trade unions, and syndicates played an increasingly significant role and appeared to be on their way to supplanting the traditional, informal channels of political process that operated alongside them. This trend was reversed after the 1950s when the governments in power adopted more authoritarian policies and suppressed formal political channels. This turn to authoritarianism coincided with the most dramatic waves of migration to Middle Eastern cities, which increased the urban populations of Egypt, Iran, and Lebanon. Thus, the rapid urbanization of Egypt and Iran since the 1950s and the absence of formal institutions connecting the general population to the state combined to provide the traditional, informal networks with an opportunity to reassert themselves.

The informal networks in Lebanon, on the other hand, followed a different trajectory into the present. They survived unscathed and were not seriously challenged by the often weak formal networks. When the formal networks did appear in Lebanon, they were simply a modern facade for the old informal patron-client and sectarian ties. Yet, regardless of the reasons for the survival of informal networks in Egypt, Iran, and Lebanon, such networks continue to provide their urban population with autonomous arenas ("zones of autonomy") where the public can address its political concerns.

Politics, however, is not the only sphere of activity for the informal networks. As Denoeux convincingly demonstrates, informal networks play an important social function by providing services that the cities do not or cannot provide. In addition, they play an essential integrative function for the uprooted, anomic, recent migrant poor of the cities. As



such, the informal networks more often than not play a stabilizing role in the life of the cities and only under certain conditions do they become vehicles of opposition to the state.

These social and integrative functions of informal networks are better demonstrated in a more detailed scrutiny of religious networks. Dividing the religious networks into two broad categories of "radical utopian" and "double edged," Denoeux argues that while the millenarian ideology of a small group of radical-utopian religious networks always places them in opposition to the state, the majority of religious networks are double edged, with a political stance that is not always oppositional. The more significant and extensive double-edged networks turn to oppositional politics only when their leadership is politically threatened by the state and then only if they have a resource base independent of the state. In the course of this discussion, the reader is also exposed to insightful suggestions about the social basis of extremist religious movements.

Denoeux draws some of the most interesting examples of the role of informal networks from the Iranian revolution of 1979 and the manner in which the Iranian clerics used informal networks to gain undisputed control over state machinery. According to Denoeux, after the revolution the clerics eliminated their rivals by using various informal networks such as the revolutionary committees. The clerics compensated their clients—the informal clusters of upwardly mobile lower- and lower-middle-class urban youth—with access to privileges and eventually with formal status. Once the clerics dominated the state, the radical clerical faction manipulated other informal networks strategically to position its clerical clients so as to undermine the standing of the better established, more prestigious, and more moderate rival clerics in the major cities.

Denoeux's account is also interesting for its synthesis of Durkheim-inspired breakdown theories of collective action with resource mobilization theory. While avoiding the pitfalls of earlier Durkheimians and devising a direct line between psychological strains of rapid urbanization and collective action, Denoeux argues that anomic recent migrants do play an important role in the urban unrest of the Middle East. If the patrons of such networks, always among the most established members of these societies, have strong incentives to engage in oppositional politics, they do so by mobilizing their clients among the recently arrived urban poor who are generally inaccessible through formal channels but who are readily absorbed into integrative informal networks. Sociologists working on collective action are well aware of the long-standing dichotomy between Durkheimian and resource mobilization theories, and Denoeux's lack of attention to this division is rather surprising. If he had acknowledged the traditional dichotomy between these viewpoints, Denoeux could claim greater credit for resolving them.

Denoeux succeeds not only in presenting a meaningful, theoretically informed, comparative study of urban unrest, but also in providing empirical case studies in which two seemingly opposing traditions of collective action are fruitfully reconciled.

*The Policy Challenge of Ethnic Diversity: Immigrant Politics in France and Switzerland.* By Patrick Ireland. Cambridge, Mass.: Harvard University Press, 1994. Pp. xiv + 327. \$52.00.

Carl-Ulrik Schierup  
*University of Umeå*

Immigrants and ethnic minorities have become distinctly articulated actors on the contemporary European political scene. They have become a political force voicing structural disjunctions and social grievances of a growing urban underclass that no responsible political establishment can afford to neglect. In a situation in which Europe is marked by a rising volume of research on international migration, however, the problem of ethnic minorities' participation in national and local politics remains a neglected issue. It appears symptomatic that one of the most remarkable studies of ethnic minorities' involvement in European politics should be done by an American. Patrick Ireland's important book on immigrant politics in two Western European countries demonstrates a discriminating understanding of the European political scene that could perhaps be credited to the privileged position of the distant observer.

Ireland's study takes the analysis of increasingly polyglot European societies beyond models, the analytical dependability of which is easily blocked by unidimensionality or conceptual abstraction, and into the realm of complex historical and institutional analysis. As such, the book represents a challenge not only to still-common primordialist conceptions of ethnicity, but also to the kind of poststructuralist conceptions of ethnic identities and ethnic politics that have put their distinctive stamp on the frontline of the 1980s and 1990s European research on ethnic relations. It does so, moreover, without taking recourse to any simplified notion of class.

Ethnicity and class are both fundamental dimensions of the immigrant presence, yet there is no direct connection linking either of them to the conspicuous political involvement of ethnic communities all over Western Europe. The structural disjunctions of advanced capitalism and the general terms on which migrants are inserted into society everywhere generate potential for protest. But the specific ways in which migrant communities phase into politics, whether employing class or ethnicity as vehicles for participation, Ireland argues, depend on the particular set of economic, political, and social conditions and on the different opportunity structure they meet in each country.

This argument is supported by an effective research design employing a two-countries/four-communities approach. A cogent presentation of the development of immigrant political representation in France and Switzerland respectively provides two distinct national patterns. In politically centralized France, immigrants soon discovered that, in order to attain any degree of power they were forced, as were other interest groups

before them, to aim for the national level. Initial restriction of the possibilities for political participation by nonnationals, and the particular character of the French left, meant that the channels of influence that were open were mainly those that earlier struggles of the indigenous working class and previous policy decisions had left available. French social welfare programs' inclusiveness of immigrants reduced the need for an extensive network of homeland-sponsored organizations. Swiss public and private welfare institutions were, in contrast, rather unresponsive to the social needs of the immigrants, which led homeland consulates and organizations to step in. This led to a higher involvement of ethnic minorities with their homelands and to the cultivation of regional or national homeland-oriented identities. This was exacerbated by the decentralized structure of Swiss society, which induced immigrants to concentrate their efforts on the subfederal level.

These national patterns are further differentiated with respect to the two local communities studied in each country. Ireland refers to the gatekeeping role of local activists in trade unions, political parties, and solidarity groups and to the adverse impact of populist movements.

The study holds wide implications for an analysis of the place of ethnic communities in European national states in crisis. Immigrants are, in an increasingly radical way, raising the "boundary problem" of who constitutes "the people," or the unit within which democratic governance is to be practiced. Governments find themselves, the author maintains, caught between growing pressures from European integration, anti-immigrant forces, employers' demands for more cheap labor, and ethnic communities' disinclination to passively accept their social and political marginalization. The French and the Swiss situations, however, illustrate two contrasting ramifications of Europe's current crisis of representation and citizenship. Swiss political institutions appear to have channeled away the immigrants' protest potential, fragmenting their political energies while substituting formal political citizenship for participation in the more limited spheres of work or school. In contrast, as they are channeled through the institution of citizenship, immigrant communities in France slowly but surely insert themselves into the French polity. Integration into the European community is currently eroding France's identity. A united Europe is, however, adopting many of the defining values of Jacobin France. Both of these changes, Ireland concludes, eventually move in a direction that adds legitimacy to the struggles of ethnic communities.

*The Political Ecology of the Modern Peasant: Calculation and Community.* By Leslie E. Anderson. Baltimore: Johns Hopkins University Press, 1994. Pp. xv + 208. \$42.50.

Marc Edelman

*Hunter College, City University of New York*

This book makes one very good point: students of peasantries have emphasized extremes of quiescence and rebellion while giving insufficient attention to the nonviolent forms of collective action that constitute the most common kinds of peasant politics. Leslie E. Anderson's study—based on 18 months of fieldwork in six villages in Costa Rica and Nicaragua—argues that each community has characteristic notions about the spectrum of acceptable political activity (from voting and petitions to land invasions and violence) and that each also has an “area of dissonance” or disagreement about the more “assertive” kinds of actions. Economic, socioeconomic, and perception-of-injustice indexes (the construction of which is outlined in an appendix) are independent variables that explain differences between and within communities. Often, poignant quotations from interviews complement the quantitative analysis. The findings that peasants employ a range of tactics and that they cooperate for political and other ends are used to argue for a broader theory of peasant politics, a theory that falls between the individualism of Samuel Popkin (*The Rational Peasant: The Political Economy of Rural Society in Vietnam* [Berkeley: University of California Press, 1979]) and the communitarianism of James Scott (*The Moral Economy of the Peasant: Subsistence and Rebellion in Southeast Asia* [New Haven, Conn.: Yale University Press, 1976]). This synthesis of rational actor and moral economy theories is termed “political ecology” because peasants allegedly view themselves as existing in interdependent relationships with their communities and the natural environment and seek to restore balance when such connections are disturbed.

Staking out a middle ground between Popkin and Scott may be long overdue, but Anderson's book hardly accomplishes this goal. Her political ecology, immodestly termed a “unique” contribution to peasant studies (pp. 119, 166), is basically warmed-over functionalism. The Nicaraguan village Pikín Guerrero, for example, “was an interactive whole whose parts all contributed to and were dependent on the others” (p. 136). Like earlier functionalisms, Anderson's paradigm is tautological: “It explains how the peasants' political ecology has led to the development of a particular ecological mentality and how that mentality leads peasants to act politically” (p. 164).

Anderson sensibly selected the villages “with an eye to studying several of the principal crops grown by peasant farmers” (p. 178). But in Pedregal, Nicaragua, she reports that they grew “a coarse, large-grain wheat” (p. 42), apparently unaware that wheat is a temperate zone crop, not a tropical crop, and that Nicaraguan peasants commonly refer to

sorghum as wheat (*trigo*, a euphemism that eases their humiliation about having to consume "animal" food). Similarly, in the humid lowland village of El Hogar, Costa Rica, Anderson incorrectly translates *yuca* (the tuber cassava) as "yucca," bringing to mind the spiky-leaved shrub found in southwestern U.S. deserts.

Mistaking the crops that her informants grew may seem trivial, but it is indicative of a broader carelessness that undermines much of the argument. There are numerous interpretative and factual errors and inconsistencies, particularly as regards Costa Rican and Nicaraguan history. In Costa Rica, for instance, all three key agricultural sector institutions that Anderson states (p. 27) were founded under "the [social democratic José] Figueres government" were actually founded under other presidents of Social Christian or conservative orientation. Moreover, Anderson's characterizations of these agencies are ahistorical caricatures. The agrarian reform institute is described as "inert" in 1980 (p. 105), even though in 1979 it acquired more land (almost 70,000 hectares) for more rural families than at any other time in its history (*Tierra con fronteras: Treinta años de política de distribución de tierras en Costa Rica*, by Isabel Román and Roy Rivera [San José: CEPAS, 1990, p. 15]). The National Production Council, a state grain board, is portrayed (p. 28) as "squeezing" peasants rather than consumers—arguably the case in the mid-1980s, when neoliberalism (and Anderson) came to Costa Rica, but not during the preceding four decades of artificially high support prices. At times, Anderson's account seems disingenuous: she reports (pp. 101–2), for example, that "a local school teacher" organized the radical peasant union UPA-GRA, even though Costa Rican scholars have described its founding by an urban cadre from an ultra-left political party. The treatment of Nicaraguan history is similarly flawed, particularly in its simplistic depiction of the Somoza regimes as maintaining social control for more than four decades solely through coercive, rather than hegemonic, means.

The generic references to "peasant society," "the peasant class," and "Latin American culture," as well as the ludicrous assertion that peasants always seek harmony with their environment, hark back to an earlier, more innocent era. If early peasant studies often romanticized village life and overlooked factionalism, social differentiation, unsustainable cultivation practices, and smallholder deforestation and encroachments onto indigenous people's lands, it is most peculiar to ignore such realities in 1980s Central America. The acknowledgments indicate that many distinguished scholars of agrarian issues commented on the manuscript. I wonder if the formal and informal processes of collegial review can no longer deeply influence work intended for a model of intellectual production and advancement that places a premium on rapid publication but discourages in-depth fieldwork and close interchange with colleagues in the countries where U.S. academics go to do research.

*Gender and Power in Rural North China.* By Ellen R. Judd. Stanford, Calif.: Stanford University Press, 1994. Pp. xiv + 295. \$37.50.

Kenneth Pomeranz  
*University of California, Irvine*

This carefully researched and provocative book touches on many important matters but is narrower than its title suggests. Essentially, it is a study of how women in Shandong province have fared in the economic reforms undertaken since 1978. The most spectacularly grim "women's problems"—female infanticide, the kidnapping and sale of women in some areas, domestic violence (including violence against women caught between the state's birth control program and the desires of their families)—are absent from the book. The book covers well the large and important subject that remains. However the narrow focus and these absences mean that the author's most sweeping claims about power and gender—particularly her claim that the importance of the family/household as a locus of women's oppression has been overemphasized by most other scholars—seem far too narrowly based.

In a variety of ways, Judd argues, "male control beyond the household is stronger than that within the household" (p. 236). This argument is partly based on a creative analysis of how state power has continued to increase in the era of decollectivization. Judd first notes the ways in which village-level governments have become more powerful, in large part thanks to the patronage they can dispense as owners of much of China's burgeoning rural industry. In this, Judd largely follows the lead of Vivienne Shue (*The Reach of the State* [Stanford, Calif.: Stanford University Press, 1988]), though she adds some interesting new examples. Meanwhile Judd also calls our attention to the less overt ways in which the state is "diffusely present" (p. 251) in villages, in large part through its power to define such crucial units of everyday life as the village and the household. It is, after all, the state that assigns people unambiguously (and often arbitrarily) to membership in one and only one household and village, thereby giving these units much firmer boundaries than they would otherwise have, and that ties many of people's most basic entitlements to their memberships in these units (rather than assigning entitlements to individuals). And in most cases, this arrangement operates to the detriment of women. Since many women work for several years in village-owned factories before they marry, but will spend their postmarriage lives as "members" of a different family and village, the state's defining power serves to cut women off from the public and private patrimonies that their youthful labor helps to build: both their wages from these years and their contribution to accumulation by village enterprises ultimately benefit their brothers' families, not their own (p. 102).

Moreover, women are multiply disadvantaged in the rural industrial sector. Management positions go almost exclusively to men (pp. 88, 97). The situation is still worse in marketing and procurement—where con-

nections essential to eventually striking off on one's own are made—in part because it is still socially unacceptable for women to undertake long-distance travel and meet with unrelated men in unsupervised settings (p. 89) and in part because women are far less likely than men to have preexisting useful connections (e.g., through military service) in other places. The contributions that women do make to village industrial enterprises seem to have little effect on the way they are seen by men—at least in one of Judd's villages even women who hold jobs away from home are referred to by the village males as "housewives" (p. 103). So while in-marrying women may benefit in a strictly material way from the surplus built up by other women—a point that Judd fails to note—they are not seen as *entitled* to a share in that surplus, as village men tend to be.

In contrast to the many negatives of the rural industrial sector, Judd gives a relatively positive account of the way women have fared in household-based rural enterprises, whether involving agricultural, commercial, or "handicraft" work. While investigators elsewhere have argued that agriculture—usually the most poorly compensated extradomestic work in the village—is being "feminized," Judd finds a more complex division of labor, in which men continue to perform many agricultural tasks (though less than women and more often work of a kind that creates or reinforces ties beyond the household; pp. 42–44, 50). And in rural "sideline" activities organized on a household basis, women often have responsibilities as important as their husbands'. Indeed, Judd argues, the degree to which the women of the household can provide "high quality" labor is "the striking distinguishing factor" between the more successful and less successful household enterprises and is the best predictor of which households the state accords the privileged status of "specialized households" (pp. 152–53, 159). In sharp contrast to the situation in village-owned enterprises, the contribution of skilled female labor to the success of household-based enterprises is commented upon by neighbors (p. 152); thus, here, women do receive recognition for their work.

Judd recognizes that household enterprises are no panacea for rural women but vigorously dissents from the assumption of many scholars that the reinforcement of the "patriarchal" rural household economy at the expense of collectives is likely to mean a step backward for gender equality: "If women's work groups outside the household have disappeared, married women show no sign of missing them. . . . Patriarchy may have lost ground in the weakening of organizational levels immediately above the household, and it often may be easier to deal with on a household basis" (p. 162).

This thesis is intriguing and—at least for the three villages surveyed—convincingly argued insofar as it relates to the economy. But it is here that one would like to know much more about a broader set of gender and power relations in these households. Are women whose economic contributions are more obvious than before necessarily more likely to have an equal voice in family planning, to be immune to violence, or

even to have an equal say in consumption decisions? For instance, many visitors to urban Chinese homes in which husband and wife made exactly the same salary have noted how far down the list of priorities a refrigerator and a washing machine—which save work mostly for women—seem to be. While some of these questions may be too intimate for surveys to illuminate them, others may not be, and at least a speculative attempt to connect women's economic roles with other aspects of their status and autonomy might have made the importance of this book evident to a larger readership. But even if Judd does not deliver all that her title promises, she has made an important contribution to our understanding of post-1978 rural China, with major implications for both women and men.

*Childerley: Nature and Morality in a Country Village.* By Michael Mayerfeld Bell. Chicago: University of Chicago Press, 1994. Pp. xi + 279. \$32.50.

Wendy Griswold  
*University of Chicago*

Like most people, social scientists used to think about "place" merely as the location or space where interesting things happened—things like interactions or GDPs or riots or patterns of inequality. Such phenomena were observed in a particular place or compared across places. Except in regional-cultural studies, place itself was usually just an intervening variable, a proxy for representing economic or other structural factors.

Not anymore. Now that discourse analysis points out that everything is a code for something else, now that postmodern theories report that place is replacing time in the organization of experience, now that human geography asserts the structuring influence of spatiality, place itself—over and above the charms or horrors of any particular place—is becoming the conceptually interesting thing. Enthusiasm for this fresh analytical angle can produce studies that are theoretically dazzling but empirically weak, of course, but that is emphatically not the case with the book under discussion here. In *Childerley: Nature and Morality in a Country Village*, Michael Mayerfeld Bell demonstrates the power this new theoretical position can have when it is used to illuminate the findings of traditional, painstaking ethnography.

Childerley is Bell's pseudonym for a Hampshire village of 475 people in southern England. It is two hours from central London, an hour from the edge of London's suburban development, and closer to the satellite cities and market towns that circle London like particles in Saturn's rings. So Childerley is, in city-planning jargon, a modern exurb of a polynucleated urban cluster as well as a largely unspoiled English village, a bedroom community as well as a centuries-old farming town, and the interpenetration of urban and rural is everywhere visible. Bell went to



this paradoxical place to study nature, "nature as a form of cultural and moral understanding." The idea that "nature" is a cultural construction has been around for a while—British historians like Keith Thomas and the growing American literature on "the wilderness" have given detailed accounts of how people have designed an ideological nature to suit their assumptions and purposes—but Bell contributes a local, microlevel dimension missing from most studies. He shows how people use a vocabulary of place, country versus city, natural versus artificial, to signify meanings that often are fundamentally social.

Bell's argument has four parts, each gracefully illustrated with scenes of Childerley life and the reflections of its residents. First, Childerley is "slightly feudal" in that it presents a two-part class structure of the moneyed people and the ordinary people, or, from the point of view of the latter, the rich and snobby versus "us." (While the ordinary people unhesitatingly describe the stratification system this way, the more affluent tend to talk about multiple gradations.) Despite sharing the same small corner of the world, and despite numerous daily contacts, these two groups seldom visit each other's homes or even the same pubs. Ordinary people have an informal, interactive, "back door" social style, while the moneyed have a more private and sedate "front door" style. Bell has great fun describing some of the differences between these styles, as when the moneyed crowd decides to straighten up the village graveyard and one of the ordinary people fumes, "Make it more tidy—so they can drive around on their tractor mowers. I was in the Fox [the scruffy pub favored by the ordinary people] not so long ago and this commuter was in there telling them about how the graveyard should be tidied up and the stones moved. And some of the locals he was talking to had their relatives buried in there! It's their suburban mind" (p. 71).

Yet despite the omnipresence of class and class distinctions, Childerleyans find the English class system with its local manifestations as morally objectionable as it is unavoidable. So they prefer to emphasize another principle of social stratification: country versus city. Both the moneyed and the ordinary earnestly believe that there is a clear distinction between rural and city folk, that the former have more community spirit and are closer to nature, and that virtue resides in the country. They have no difficulty categorizing other people as being "not really country" (e.g., the disdainful reference to the "commuter" above) versus being "a true countryman, born and bred." And—this is the second part of Bell's thesis—"by thinking and talking country, Childerleyans think and talk class—and often avoid the moral ambiguity of the latter" (p. 105).

There is more to the country/city distinction than a veiled expression of class prejudices, however. Childerleyans hold nature as an ideal and, much as they admit the difficulty of determining just where "nature" ends and something else (artifice, culture, city) begins, they strongly believe in a pastoral gradient wherein some things and some people are closer to nature than others. And why is closeness to nature seen as desirable? As the third part of his thesis, Bell makes the intriguing case

that just as George Herbert Mead's "generalized other" constitutes a social referent through which people construct their identities (the "me" in Mead's model), so do Childerleyans postulate a "naturalized other" from which they construct a "natural me." The virtue of the natural other is precisely that it is *not* social, that it is not contaminated with the pursuit of interest that is taken to be part and parcel of human social life. Thus as an alternative to their participation in a collective conscience tainted with materialism and self-seeking, Childerleyans resort to a "natural conscience" formed in interaction with nature, and this latter conscience is one with which they feel most authentically at home. As one man put it, "I think it [the country] provides for me a kind of secure emotional base, literally. Just my relationship with the physical countryside. . . . The only thing which is constant is my own relationship with the physical environment, as opposed to the social environment" (p. 152). Childerleyans of both classes assert the same idea of what Bell summarizes as "a truer *me*, a *me* apart from social interests," even though the former sometimes is compromised because of the demands of the latter.

The inevitable question arises of whether this natural conscience, ostensibly orthogonal to class divisions, is in fact simply a reflection of them. The answer to this, which constitutes the fourth part of Bell's exploration, is not simple. It seems to be the case that people's experience of nature does indeed vary along class lines and according to class styles. For example, the moneyed have a formal and privatized appreciation of the country, as when one of them complains about more and more people invading the country and disturbing its (and his) isolation: "For years I've walked this countryside. And it's magic to go out into the countryside where you won't meet anybody." On the other hand, ordinary villagers never mention solitude and describe nature as something they are in direct interaction with, like the woman with severe medical problems who rejects her doctor's advice to move into town because she can't bear to leave her friends: "I spend hours in the kitchen watching my birds. And you know, they're so cheeky, they don't fly away when I go and put the peanuts out. They perch right there, looking at me right close, I could pick them, pick them off the branch" (p. 175). Similarly, both classes value gardens, but while the moneyed crowd fashion secluded gardens behind their houses with an eye to aesthetics, the ordinary folks establish open gardens in the front filled with a hodgepodge of flowers, lawn ornaments, domestic animals, and sites of various activities like working on the car. In a somewhat different vein, virtually all ordinary Childerleyans and many moneyed ones strongly disapprove of fox hunting, but for different reasons: the latter scorn city people dressed in fancy new habits and riding expensive horses for trying to buy themselves into being "country," while the former voice egalitarian objections to the essential unfairness of the chase. So the experience of nature itself can be seen as a social projection.

But not entirely, according to Bell, who rejects the simplistic view that participation in a natural conscience merely reflects one's social position.

Instead he draws on the image of Childerley's church bells to speak of resonance, his metaphor for the observation that "we tend to think a thing more true when we can find it in a range of experience" (p. 234). Thus when people experience nature they seek points of resemblance to their social experience: a mother of a large family may see a nurturing, benevolent nature while a small businessman may see it in competitive, Darwinian terms. The reverse is also the case, whereby natural conscience shapes social experience through partial resemblance: the woman who sees the birds as her friends and equals has no trouble standing up to—being "cheeky" to—her doctor. Like most sociologists, Bell suspects social experience is the "tenor bell" in the chain of resonation (he briefly mentions other experiential possibilities such as the body, offers a rather superficial discussion of gender as just one aspect of the social, and ignores other contenders such as language and transcendence), but he wisely abjures determinism: "If a way of hearing and experiencing suits different realms, that way seems intuitively and ideologically more sensible to us. Indeed, it would be surprising if this were not so. But the direction of how that agreement comes to be—from the generalized other to the natural other, from the natural other to the generalized, or from an interplay between these and other realms of thought and experience—is not significant to us. As long as we get there" (p. 239).

What Bell calls resonance has much in common with elective affinity and with the formal properties of metaphor, and I believe this view of how people relate the different domains of their experience is absolutely right. Rarely has this insight been applied to such different levels of human action and cognition, from the back-door interactions to the stratification system, from one's sense of self to one's participation in the cosmos. The reader, this reader anyway, finishes *Childerley* with the feeling that she has just returned from visiting a remote Hampshire village and has learned something, not just about that place, but about human social life lived in other places and lived through place itself. Bell has succeeded in doing what all ethnographies aim at, which is to convey something of the experience of T. S. Eliot's traveler in "Little Gidding": "We shall not cease from exploration/And the end of all our exploring/Will be to arrive where we started/And know the place for the first time."

*Demography of Aging*. Edited by Linda G. Martin and Samuel H. Preston. Washington, D.C.: National Academy Press, 1994. Pp. x+411. \$39.00.

Işık A. Aytac  
*University of Chicago*

In the introduction to this collection of essays, Samuel H. Preston and Linda G. Martin set the stage for examining how "the demography of aging" has emerged as a field over the last decade as a result of the

significant demographic changes of the 65 and older age group. Referring to United Nations' projections, they argue that population aging is becoming "one of the most important social phenomena of the next century" (p. 3), with serious implications for the quality of life depending on how social institutions respond to this challenge.

In the first essay, Ronald D. Lee develops an accounting framework and uses it to show how systems of transfers are affected by changing population age distributions. He provides a comprehensive discussion of the links between fertility and mortality, population age distributions, and interage transfers through the family, public sector, and financial markets.

The next two chapters address the economic dimensions of the demography of aging. First, Joseph Quinn and Richard Burkhauser discuss how retirement has become a phase-out process, rather than an abrupt exit from the labor force. Next, Douglas Holtz-Eakin and Timothy Smeeding focus on the economic circumstances of the aged and on the motives for intergenerational transfers. They emphasize the importance of distinguishing among income, wealth, bequests, and inter vivos transfers. They also stress that the American elderly cannot be treated as a homogenous group; while married elderly are the wealthiest, elderly single women are the poorest in comparison to other countries. Both chapters provide a wealth of information on existing data, especially the new Health and Retirement Survey and AHEAD surveys.

The next two chapters focus on kin availability and care of the elderly. Douglas Wolf discusses the demographic changes in kin availability and proximity to elderly, and how these changes affect elderly living arrangements on individual, national, and crossnational levels. This is followed by Beth Soldo and Vicki Freedman's work on the division of labor among the family, the market, and the state in support and care of the elderly. They argue for the need to understand the degree to which individuals substitute the personal care they give to family with financial support or paid workers and to what extent increases in state support lead to decreased family care.

In the next chapter, Kenneth Manton and Eric Stallard discuss how, with the aging of American society and the high costs of health care, medical demography has reemerged as a field with significant policy implications. They point out that, since the elderly are more likely than younger persons to suffer from multiple diseases, elderly health care demands are different from those of younger people and require more complex models to study.

In chapter 8, Samuel Preston and Paul Taubman examine class differences in adult mortality and health using two approaches—economic and social psychological. They review recent evidence on how education, income, and occupational status are associated with health, mortality, and other health-related behavior (i.e., social relations and social psychological factors). They also point out that black-white differences in mor-

tality and health are primarily due to racial inequality in income and education.

With the elderly (65 and older) expected to represent around 20% of the population by 2030, Frank Bean, George Myers, Jacqueline Angel, and Omer Galle examine the geographic distribution of the elderly and how it has changed over time in the United States. They discuss the role of migration and immigration in the spatial distribution and concentration of the elderly and conclude the chapter with international comparisons, knowledge gaps, and data needs in the field.

In the final chapter, Linda Martin and Kevin Kinsella provide a review of the research on the demography of aging in developing countries. Research in developing countries, which began in the 1980s, covers substantive topics in basic demography, mortality and health, family demography, population distribution and migration, and economic activity and well-being. Martin and Kinsella end their chapter with a discussion of data needs. They also provide a useful list of existing data by country and region in the appendix.

This edited volume brings, from a variety of disciplines, experts who address labor force participation and retirement, economic well-being, social support by family, disability, mortality and health care, and aging in developing countries. The book contains a lucid and useful review of previous research, data sources, data needs, and recent research that has significant policy implications. It is a first in the emerging field of demography of aging and is highly recommended for scholars, policy-makers, and advanced students.

*Old Age: Constructions and Deconstructions.* By Haim Hazan. Cambridge: Cambridge University Press, 1994. Pp. xi + 126. \$44.95 (cloth); \$12.95 (paper).

Anne Foner  
*Rutgers University*

According to Haim Hazan, most research on old people is flawed, ignoring older people as subjects and failing to understand ways in which knowledge about aging is produced. In offering a corrective, his book critically examines approaches in the sociology of age and provides an alternative perspective.

Discussing languages about old age, Hazan argues that current usages perpetuate misunderstandings about the aged. First, the old are represented as distinct from and alien to society. They are assumed to be nonproductive, and aging itself is seen as a disease. These views have implications for the way the old are treated and the way the old react.

Second, pervasive stereotypes of the aged are not only misleading but camouflage the social arrangements for old people. For example, the

image of older people as incapable of initiative and sound judgment about their own needs provides a rationale for denying independence to older people who enter old age homes.

Taking the viewpoint of older people themselves, Hazan argues that older people use various strategies to counter these false images: ritualized conformity to social expectations, withdrawal, rebellion, taking up a new occupation, travel, joining a political movement. These various strategies parallel Robert Merton's analysis of modes of adaptation to the disjunction between social goals and institutionalized means for achieving these goals (*Social Theory and Social Structure* [Glencoe, Ill.: Free Press, 1957]). It is puzzling that Merton's model is not mentioned in the discussion of these various strategies but is only cited in a later chapter.

Turning to what he calls "the theoretical trap: the missing language" (p. 39), Hazan asserts that the major theories of old age are unsatisfactory and that social scientific language is inadequate for understanding "the parameters of signification among the old" (p. 50). The concept of role comes in for particular criticism. Hazan argues that it is inappropriate as a focus of analysis, description, and understanding of the world of older people, because there are too few meaningful roles for older people. He criticizes even volunteer roles, since they are devoid of status and financial rewards and represent a mock inclusion in society. One could argue to the contrary: it is not the concept of role that deprives older people of roles; indeed the concept helps us to highlight such deprivations.

In Hazan's own thinking two basic elements are crucial for understanding old age: control and meaning. Concerning control, Hazan asserts that age is used as one of the most commonly accepted criteria for social evaluation and location, with age boundaries enforced by both formal and informal mechanisms. Age criteria influence everyday relationships, situations, and discourses, although the importance of age varies with type of relationship and at different periods of the life course.

Concerning meaning, Hazan touches on cultural boundaries defining old age. He maintains that the aged are perceived as located between life and death. The aged function for the rest of society as a buffer between life and death. As such they have been stripped of their roles and excluded from the category of full human beings. They are in a permanent state of limbo. Such perceptions underlie negative attitudes about the aged.

But understanding both structural and cultural boundaries does not get at the experience of being old, according to Hazan. Hazan argues that to understand old age scholars must explore "their" culture, not "our theories," and try to grasp the ambivalent nature of the inner and outer experiences of the old. Just how we are to go about this endeavor will not be clear to most readers unfamiliar with much of the specialized language used here.

Hazan's critical analysis is provocative. But whether his assertions about social attitudes to old age in modern society are in accord with

empirical reality is an open question. For example, is aging identified with ugliness, evil, and horror? Is there reluctance to have physical contact with old people? Is old age generally perceived as a buffer between life and death? Is fear of death strongest among the middle-aged? (See pp. 68, 72; for a rounded review of theory and research on death and dying specifically, see John W. Riley, Jr., "Dying and the Meanings of Death: Sociological Inquiries" [*Annual Review of Sociology*, vol. 9, 1983].) To be sure, the very assumption of Hazan's book, that the current tools and concepts of social science cannot grasp the full realities of old age, challenge such questioning.

Hazan himself has conducted ethnographic research. But it is not clear from the several empirical illustrations he provides just what tools he uses. It is noteworthy that only one of what he calls "ethnographic reflections," a case study based on a self-help organization, the University of the Third Age in Cambridge, England, points to the potential for creativity among the old. The others—describing members of an Israeli old age home, a day center for elderly Jewish residents in London, Israeli media presentations on the plight of the elderly, and a brief first-person account of the experience of being old—hardly underscore the positive or represent diversity in the life situations of the old.

The book, with its obscure language, will not be accessible to all readers. But it raises issues for social science theorists and researchers studying old age to ponder.

*Succeeding Generations: On the Effects of Investments in Children.* By Robert H. Haveman and Barbara L. Wolfe. New York: Russell Sage Foundation, 1994. Pp. xi + 331. \$34.95.

Susan E. Mayer  
*University of Chicago*

The U.S. government's official poverty statistics show that in 1994, as in each of the past five years, the poverty rate among children increased. This and other indicators of the ever-worsening circumstances in which children live have created renewed interest in what government and families can do to help children. *Succeeding Generations: On the Effect of Investments in Children* by Robert Haveman and Barbara Wolfe is one such effort. These authors try to determine which characteristics of families, neighborhoods, and social policies increase children's educational attainment and labor force activity and decrease teenage out-of-wedlock childbearing and subsequent welfare receipt.

To do this they adopt what they call an "investments-in-children" framework. This framework follows the lead of Gary Becker and his colleagues, who argue that children's success is a function of their biological and cultural endowments and the investments that parents and society make in them. Like other economists, Haveman and Wolfe view

investments as the result of a maximization process in which parents and government choose how to allocate the scarce resources of time and money. Similarly, children themselves make decisions regarding their education, work, and fertility that may maximize their well-being. Although the maximization model assumes choice as a heuristic device, Haveman and Wolfe make choice the centerpiece of their framework.

Many sociologists will find the emphasis on choice irritating. Haveman and Wolfe write, "Parents choose the sort of monitoring, disciplinary, nurturing, and expectational environment that they create for their children" (p. 32). But some readers will think that very poor families have so few alternative environments from which to choose that they have no "real" choice at all. Of course, this is an empirical question: Is there a lot of variation in the environments of poor children? But it is not the sort of empirical question that Haveman and Wolfe try to answer. Instead they model the effects of a compilation of family background and neighborhood characteristics that other researchers have found to influence children's outcomes. Like other researchers, they lack direct measures of both the decision-making process and the actual investments that parents make in their children.

The authors use the Michigan Panel Study of Income Dynamics to produce ever more complex estimates of the family and neighborhood factors that affect children's success. They begin with zero-order correlations between each outcome and race, parental education, indicators of family stress, parental income, neighborhood attributes, and family circumstances including religion and the number of years the family head reported being disabled.

They then produce what might be considered a reduced-form estimate of the influence of many of these factors on each outcome. The next model includes these factors as well as neighborhood attributes and measures of family stress, including the number of years the head was unemployed, the number of changes in parental marital status, the number of household location moves, and whether the family ever received welfare. The subsequent chapters estimate models specific to each outcome. Each chapter includes a set of useful simulations showing, for instance, the probability of high school graduation conditional on parental education or on the number of family moves.

Readers will disagree about the adequacy of Haveman and Wolfe's estimates. For instance, they include no measure of either parents' or children's cognitive abilities. Since this sort of endowment is crucial in Becker's model, this may be a serious limitation to their results. Instead of parental income, they usually estimate the effect of the number of years a family lived in poverty. This tells us the difference between often being poor and never being poor, but not the difference between living in a family whose income is half the median income compared to one whose income is twice the median.

Haveman and Wolfe draw their conclusions from their more elaborate models, not their reduced-form models. This may not always be appro-



priate. For instance, they conclude that the effect of parental income is qualitatively small on all outcomes. This conclusion follows from equations that control neighborhood composition and parents' marital status. Since parental income is an important determinant of neighborhood composition and at least partly a determinant of divorce, their estimates may understate the overall effect of eliminating poverty.

In the end Haveman and Wolfe's conclusions and policy recommendations are much the same as those of everyone else. Parental education and marital status are important, children who face a "large number of adverse parental choices" (p. 262) fare worse than those who face none, stressful family circumstances hurt children, as does living in a bad neighborhood, and so on.

Haveman and Wolfe recommend that policy foster more education for children, increase the income available to families, foster intact families, improve living conditions in poor neighborhoods, and provide support or counseling for children who experience stressful circumstances. Most people will not quibble with these goals, although they might disagree about how to achieve them.

*Beyond Adolescence: Problem Behavior and Young Adult Development.*

By Richard Jessor, John Edward Donovan, and Frances Marie Costa. New York: Cambridge University Press, 1991. Pp. 312. \$39.95.

Robert B. Cairns

*University of North Carolina at Chapel Hill*

Understanding the integrity of the life course, tracing its continuity over large segments of time, distinguishing what is ephemeral from what is lasting, grasping the role that the past plays in shaping the future—all these, and more, are issues that yield only to research that is longitudinal and developmental in design.

(Jessor, Donovan, and Costa 1991, p. 3)

In the mid-1980s, several new longitudinal investigations were launched, stimulated in part by the findings of R. Jessor and S. L. Jessor (*Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth* [New York: Academic, 1977]) that the problems of adolescence occurred in packages rather than as isolated characteristics. In the mid-1990s, investigators are being challenged to demonstrate that the new longitudinal analyses are worth their cost in time and dollars. *Beyond Adolescence* may be used—as was the earlier monograph—to justify the logic and continued funding of longitudinal research on problem behaviors and how they develop.

Taking up where the earlier volume left off, *Beyond Adolescence* reports continuities and changes in patterns of substance use, deviance, and careers by following into adulthood the adolescent subjects who were

earlier studied. There is ample evidence for continuity. As expected, heavy drinkers in adolescence are likely to be heavy drinkers in adulthood, and illicit drug users in adolescence tend to become illicit drug users as adults. However, there are some twists in the story. One finding is the developmental trend toward increasing conventionality among subjects in their late twenties, especially among persons who were most deviant in high school and college. These subjects report decreasing levels of illicit drug use and fewer episodes of heavy drinking. The rehabilitative effects of maturity—hinted at by cross-sectional national crime statistics and found in the longitudinal studies of Werner in Kauai and Farrington in London—are confirmed in the present data set. In terms of the implications for theories of deviance and public policy, this age-related rehabilitation effect deserves close attention.

Jessor, Donovan, and Costa conclude, "The data support the notion of a syndrome of problem behavior in young adulthood, a notion that had previously been shown to hold for adolescence" (p. 126). How does this statement jibe with their own findings that occupational success and educational attainment in young adults were not jeopardized by the syndrome of problem behavior? Part of the answer lies in how the problem behavior category was generated. As constructed in *Beyond Adolescence*, the adult problem behavior syndrome includes four factors of substance abuse (alcohol, marijuana, other drugs, and cigarettes) and a poorly defined "general deviance" factor. The coherence of the problem behavior syndrome in adults seems determined in large part by the investigators' decisions on which variables to enter into the LISREL solution and which to omit.

The authors' proposal that there are correlated problems in youths and adults challenges traditions in psychiatry and psychology that have balkanized mental disorders into separate and unrelated categories. The concepts of "problem behavior syndrome" and "co-morbidity" question the realpolitik view that society can stamp out unhealthy adolescent behaviors one by one. In the 1990s, the one problem—one intervention model continues to support separate national programs for preventing violence, drug abuse, dropping out of school, and teenage parenthood. *Beyond Adolescence* reasserts the need to view these programs in the context of the integrated nature of social and personality development. It also implicitly raises the question, Why attempt to change adolescent problem behaviors if the problems are self-correcting?

Since the procedures of *Beyond Adolescence* were designed in the early 1970s, how well do they compare to today's standards? On many counts, Jessor and his team were ahead of their time. The research was thoughtfully designed and carefully executed, with 89% recovery of the original sample of 632 subjects over a nine-year interval. Despite the pitfalls of data tyranny, the information was painstakingly analyzed and clearly reported.

Although measures in social science age slowly, they are not timeless. Over the past two decades, the longitudinal measurement strategy of

choice is shifting from the self-report survey questions employed by Jessor and his coauthors to multiple-agent, multiple-method assessments. Recent longitudinal data indicate that, given the chance, most people construct balanced stories about themselves and their current circumstances. Relative to other measures, self-reports provide (a) complementary inter-correlations within domains, (b) smoother growth curves over time, and (c) enhanced versions of the self, with most persons reporting themselves above average. In the present research, it is likely that the concrete outcome variables reported upon (e.g., job success, educational achievement, and substance use) are captured reasonably well by self-reports. But there is a need to compare the results with other longitudinal investigations that employ information sources beyond the self.

As noted above, *Beyond Adolescence* reports few reliable associations between adolescent problem behavior and adult occupational achievement or educational attainment. In contrast, other recent studies, including D. Magnusson (*Individual Development from an Interactional Perspective* [Hillsdale, N.J.: Erlbaum, 1988]) and L. Pulkkinen ("Continuities in Aggressive Behavior from Childhood to Adulthood," *Aggressive Behavior* 19 [1993]: 249–63), find that problems of adolescence have sequelae in multiple domains of adult adjustment. They find that the serious problems of adolescence are not always self-correcting.

Why the discrepancy? Or is it really a discrepancy? The finding on the lack of spillover may reflect the tendency for subjects to provide balanced stories of their lives. It is also the case that the protected futures of deviant youths in *Beyond Adolescence* may reflect the relatively privileged samples studied by Jessor and his coauthors. Virtually all members in their university sample were white (98% of the men, 95% of the women), college graduates, and from middle-class (or higher) socioeconomic origins. The same holds for the members in the high school sample, where high school dropouts had been eliminated by sample selection and the school served "a community that is the site of a major university with a national reputation for its research" (p. 50). Such levels of academic achievement, community placement, or family status could mitigate against continuing problems or deal with them when they occurred.

On the other hand, why not study the development of persons-in-context? In this regard, Jessor and his colleagues have been careful in describing the special characteristics of their sample. Their work suggests that the developmental trajectories of problem behavior may be relative to the context in which they are observed. Heavy substance use may have markedly different consequences for adult adaptation depending on whether persons grow up, say, in the advantaged neighborhoods of Boulder or the impoverished ones of inner-city Denver. On this count, the longitudinal study of well-characterized local samples can provide information on developmental pathways that may be obscured in national probability studies.

In sum, *Beyond Adolescence* demands careful study. Its predecessor (Jessor and Jessor 1977) helped launch a new wave of longitudinal re-

search, and this volume can help justify its continuance. The work also demonstrates how rapidly the methods and analyses of longitudinal research have advanced in the last two decades. Despite these hazards, Jessor, Donovan, and Costa were doubtless correct when they argued that the basic issues of the human life course will "yield only to research that is longitudinal and developmental in design."

*Parents' Jobs and Children's Lives.* By Toby L. Parcel and Elizabeth G. Menaghan. New York: Aldine de Gruyter, 1994. Pp. xiv + 214. \$39.95 (cloth); \$18.95 (paper).

Janet Currie

*University of California, Los Angeles*

Over the past 30 years, a social revolution has taken place—in 1964, 77% of women with children under six years old stayed home; today, only 40% do so. In this thoughtful and ambitious book, Toby Parcel and Elizabeth Menaghan ask, What implications is this revolution likely to have for children's lives?

The book is innovative in many respects. First, while previous authors focus only on the effects of maternal employment, Parcel and Menaghan treat mothers and fathers symmetrically. For example, they ask whether changes in fathers' work patterns can offset changes in mothers' employment patterns, a question that is particularly intriguing in light of the continuous decrease in men's labor force participation over the last several decades. Second, they stress job content, as measured by "occupational complexity," as well as wages and work hours. Third, they use longitudinal data from the National Longitudinal Surveys' 1986 and 1988 child-mother data (NLSCM) to examine the effects of both past and current conditions on children's welfare.

In addition, the authors attempt to integrate and test the theoretical insights from three different perspectives. They examine Melvin Kohn's theories about the relationship between the nature of parental employment and the extent to which children are taught to be self-directed and to internalize behavioral norms, the theories of James Coleman regarding the importance of "family social capital," and Gary Becker's theory of the family, with its emphasis on utility-maximizing behavior.

Parcel and Menaghan use the NLSCM data to examine the effects of parental employment on home environment, children's cognitive outcomes, and incidence of behavior problems. They also examine the determinants of parental employment patterns themselves. They find that the relationships between parental employment and children's lives are complex and defy simple generalization. However, the authors draw four broad conclusions. First, "parental resources blunt the adverse effects of difficult work and family conditions." Second, "the effects of workplace

conditions vary depending on parental family circumstances." Third, "the effects of one-parent's working conditions vary depending on the working conditions of the other parent." And fourth, "occupational quality and employment status effects are also interactive" (p. 158).

If the book has a weakness, it is that the observed relationships between parental employment and child outcomes often seem to be interpreted in a causal way. For example, the authors find that a summary measure of the home environment is lower for children whose mothers work in jobs with low occupational complexity. From this they conclude that "mothers who work in occupations with more substantively complex activities create home environments that are more cognitively enriched and more affectively and physically appropriate than those who work in occupations with less complex activities. . . . We have begun to trace one facet of the causal chain that links parental background with child outcomes" (p. 61). It is likely, however, that the mother who is able to find an occupationally complex job is also able to provide an environment rich in the characteristics rated by the home-environment scale whether or not she is employed. That is, the characteristics of the job may act as proxies for unobserved characteristics of the mother rather than for any true effect of "complexity." The fact that dummy variables for race and Mexican ethnicity have such large effects in the same regression suggests to me that omitted variables are important. I would not attempt to interpret those coefficients as links in any "causal chain." It would be possible to go further with these data than Parcel and Menaghan have and to estimate models that include fixed effects to capture all the fixed characteristics of the household. Perhaps we will see these estimates in future research.

There are other decisions readers may query. First, the analysis focuses on the sample of children who were between three and six years old in 1986 and whose mothers were currently employed. The inclusion of mothers who were not currently employed would have increased the sample size and could possibly have added useful variation. Second, in families with more than one eligible child, only the youngest is included in the sample; hence, potentially valuable variation within families is ignored. Third, missing values are systematically replaced with variable means—a decision that may also have reduced the variance in variables of interest. Fourth, readers unfamiliar with the NLSCM might have appreciated a fuller discussion of the child outcome measures that are used and of whether the estimated effects are big enough to be empirically important (whether or not they are statistically significant).

The bottom line, however, is that Parcel and Menaghan deserve credit for single-handedly broadening the scope of the debate on this important question. Their results will doubtless inform policy discussions and set a rich agenda for future research as the 1990, 1992, and 1994 waves of the NLSCM are analyzed.

*Flesh Peddlers and Warm Bodies: The Temporary Help Industry and Its Workers.* By Robert E. Parker. New Brunswick, N.J.: Rutgers University Press, 1994. Pp. xi + 174. \$40.00 (cloth); \$14.00 (paper).

Alison Davis-Blake

*The University of Texas at Austin*

According to Robert Parker, the purpose of this book is "to examine critically the emergence and growth of the temporary workforce" (p. 1). Although this statement implies that Parker will approach the issue of temporary work from a broad perspective, the scope of his book is constrained in two reasonable but important ways. First, the book focuses on the consequences of temporary work for temporary workers. A discussion of the social and political consequences of the dramatic increase in temporary workers flows naturally from the rest of Parker's work and would have been a welcome addition to the book. Second, the book focuses almost exclusively on temporary workers who are employed in peripheral jobs. Parker gives little attention to the phenomenon of increasing use of temporary workers in core jobs.

The book explores three basic topics. Chapters 1 and 3 provide a description of the temporary workforce based on a review of existing literature. These chapters offer a rich and detailed picture of the variety of forms of temporary work ranging from labor pools composed largely of homeless individuals, temporary farm workers and "guest workers," employees of temporary help firms, employees of in-house temporary agencies, and the growing number of temporary faculty and government employees. These chapters are notable for the breadth and detail in which they describe contingent work; however, they do not attempt to document the size of the various components of the contingent workforce. The chapters also sometimes imply that temporary work is the primary form of contingent work. Although Parker briefly discusses part-time work, other common forms of contingent work are unexplored (e.g., employee leasing, independent contracting).

Chapters 4 and 5 document the impact of temporary work on the lives of temporary workers. These chapters are based on data collected by Parker (interviews with 40 employees of temporary help agencies and four months of participant observation as a temporary worker). These chapters are an important contribution to current knowledge about temporary workers. While some of the reports in these chapters are not surprising (e.g., temporary workers toil for relatively low wages and no benefits), two of Parker's findings go well beyond what has been reported elsewhere. First, Parker does an excellent job of portraying the high level of uncertainty that pervades every aspect of a temporary worker's life. Parker notes that temporary workers experience uncertainty not only about the timing and duration of their employment but also about many mundane matters that permanent employees take for granted (e.g., how to locate their place of employment and plan for commuting and parking,

the nature and conditions of their work including issues such as the number of breaks, the availability of food and beverages, and the safety of the workplace). Second, Parker also describes in detail the tense interactions between permanent and temporary employees (relations between the two groups are at best polite but cold; at worst, permanent workers treat temporary employees as second-class citizens). Parker's work is a nice addition to Vicki Smith's (*Work and Occupations* 21 [1994]: 284–307) research that documented similar kinds of hostilities from the viewpoint of permanent employees. Although these chapters could have been improved by a more rigorous analytic approach, they provide a rich foundation for future research on the consequences of temporary work.

Chapters 2 and 6 discuss the firms that hire temporary employees. Chapter 2 describes temporary help agencies, and chapter 6 discusses the firms on whose premises temporary workers are employed. These chapters are based on a review of the existing literature and on Parker's interviews with 11 branch managers of temporary help agencies. Chapter 2 documents the explosive growth of the temporary help industry and then discusses employee selection practices in the industry. Parker reports that temporary help agencies use highly bureaucratic selection procedures to weed out "undesirable" applicants, thus increasing the agencies' ability to provide a standard "product" and to reduce uncertainty for their clients. This chapter's insights into the similarities between temporary help service firms and other bureaucracies are valuable. The chapter would have benefited by some discussion of the extent to which the hiring process for temporary workers differs from the hiring process for permanent employees. The literature on organizational hiring standards (see, e.g., Yinon Cohen and Jeffrey Pfeffer in *Administrative Science Quarterly* 31 [1986]: 1–24) suggests many similarities between the two processes. Chapter 6 is a useful summary of arguments made by Richard Belous, Barry Bluestone, and Bennett Harrison, among others. Parker notes that increased use of temporary workers is an important way in which firms gain labor force flexibility and cope with increased economic uncertainty.

This book is a good introduction to the nature of temporary work and to the variety of issues faced by typical temporary workers. Researchers and policymakers interested in the effects of contingent work on temporary employees will find this book particularly useful.

*The Contradictory College: The Conflicting Origins, Impacts, and Futures of the Community College.* By Kevin J. Dougherty. Albany: State University of New York Press, 1994. Pp. xiv + 365. \$21.95.

Walter R. Allen  
*University of California, Los Angeles*

This book alternately provokes, educates, and disappoints. Kevin Dougherty sets as his task to "resolve the debate between the critics and

defenders of the community college. . . . What is the impact of community college? . . . Why and how did the community college develop? . . . Why did the community college differentiate its originally academic program to become so strongly vocational?" (p. 7). To accomplish this task the author marshals impressive data from historical, institutional, and other published sources.

The book is organized into five major sections. It begins with an overview that frames ongoing debates about community colleges. These debates regard the role of community colleges in expanding access to higher education. Critiques from the perspectives of pluralist functionalism, instrumental Marxism, institutional theory, and the autonomous state are presented systematically. Dougherty then summarizes (some would say hypothesizes) how advocates of each perspective would answer the three pivotal questions about community colleges posed in this book (those of impact, origins, and vocationalization).

This approach of presenting competing perspectives on specific questions, while innovative, sometimes becomes tedious, disruptive, and difficult to follow. Readers are confronted with long comparisons of competing perspectives on various questions. Ultimately this organizational device serves more to confuse than to enlighten, and in many respects it becomes needlessly repetitive. It would have been better to have outlined the assumptions and analysis of a particular perspective and then to have left the task of applying this perspective to the reader's imagination.

Section 2 focuses on outcomes in an attempt to assess the impact of community colleges on students, the economy, and universities. What emerges is a clearer statement of the contradictory tendencies of community colleges. These institutions protect elite state universities from the pressures of mass aspirations even as they provide opportunities for otherwise excluded students. Ultimately, however, community colleges are shown to debilitate, rather than facilitate, the efforts of the overwhelming majority of their students attempting to earn B.A. degrees. What is most interesting in the book's treatment of these issues is its demonstration of the link between competing, often antithetical goals and unsuccessful student outcomes. Certain features of community college organization and the population those colleges serve make it more difficult for students to persist and to successfully transfer to and survive in a four-year college.

The larger portion of the book is devoted to sections 3 and 4 which consider the origins and expansion of community colleges and the emerging emphasis on occupational education. It is on these points that the book shines. Readers are presented with a thorough consideration of the roots and evolution of community colleges. The book offers a wealth of interesting, informative, and insightful information on the history of community colleges. These sections show the often unsystematic, uncoordinated manner in which these institutions evolved. Decisions by government officials, college administrators, students, and private industry leaders combine to shape the reality of these schools. Placing these deci-



sions in a larger social, political, and historical context helps debunk notions of conspiratorial master plans. At the same time, readers are given insight and information that reveal undeniable consistency in the decisions of multiple actors, decisions that produce similar results—the overwhelming number of students whose desire to earn B.A. degrees is undermined by attendance at community colleges. This point is powerfully made in the book's comparison of the origins and evolution of community college systems in New York, Illinois, California, and Washington. Although conceived and nurtured in different sociopolitical contexts (as defined by state control, student demography, and economic base), these systems produce comparable outcomes.

The book returns in its conclusion to a point made at the beginning: community colleges are multifaceted, contradictory institutions. Thus, attempts to interpret the roles that these institutions play in higher education must be undertaken carefully. Without doubt, community colleges are major actors in the drama of U.S. higher education. In 1991 nearly six million students were enrolled in two-year colleges, as compared to the roughly eight million students attending four-year schools. The students attending community college in any given year are disproportionately nontraditional—that is, persons of color, females, and students from lower-class backgrounds. Fewer than one-fifth of these students transfer into four-year schools, and fewer still go on to obtain B.A. degrees.

The mix of students attending community colleges for remedial learning, vocational training, and prebaccalaureate study varies across regions and across schools within regions. There is also considerable variation in the relationship between these schools and key community actors—high schools, employers, and government officials. These factors ultimately shape the character of (and respond to) the college in question. To the degree that this book provides a lens for the systematic consideration of community college origins, evolution, and outcomes, it provides an important service. Perhaps this book's most important contribution is its attempt to confront the inherent dilemma of a society whose rhetoric of meritocracy and democracy must somehow be reconciled with the reality of blocked opportunities and structural inequality.

*Neighbourhood Organizations and the Welfare State.* By Shlomo Hasson and David Ley. Toronto: University of Toronto Press, 1994. Pp. ix + 387. \$65.00 (cloth); \$24.95 (paper).

Brian K. Gran  
*Northwestern University*

In their book, *Neighbourhood Organizations and the Welfare State*, Shlomo Hasson and David Ley and coauthors (in chaps. 2, 3, and 4) show how welfare state development and longitudinal changes in local

government influence neighborhood organizations. They relate various theories of collective action to social change while describing neighborhood organizations in Vancouver and Jerusalem.

Why compare Vancouver and Jerusalem? The authors point out that both cities are part of young, democratic states with mixed economies and important neighbors. Vancouver and Jerusalem have large immigrant communities and contend with the dilemma of fiscal restriction and increasing social welfare costs. This last tension, between economic development and social welfare, is what the neighborhood organizations are struggling against.

Hasson, Ley, and coauthors make three important contributions to social science research. The first contribution is their description of how local residents mobilize into collective action in the context of the urban situation in which they live. They develop four types of neighborhood organizations: ratepayers' associations, paternalistic associations, protest organizations, and those that cooperate with the state. Hasson and Ley et al. contrast failures and successes of each kind of group in Vancouver and Jerusalem. They account for the role of physical, historical, and socioeconomic circumstances and changes and human agency in their influences on the four types of neighborhood organization. They note that over time the goals of neighborhood groups have shifted from material to "postmaterial" territories. For example, while some neighborhood groups sought housing for residents, others demanded political participation.

The second contribution Hasson and Ley et al. make is their description of exchanges between neighborhood organizations and different levels of government. They relate interchanges between local and higher levels of the state; the local, provincial, and federal governments in Canada; and the local and central governments in Israel. An important part of this contribution is their examination and application of the concept of neocorporatism. In their analysis, neocorporatism is the cooperation between local neighborhoods, the state, and the private sector. Hasson and Ley et al. contrast neocorporatism with the co-optation of a neighborhood group by a local government. They conclude that researchers should reconceive corporatism as having conservative, liberal, and pluralist forms. The authors suggest that liberal corporatism may be used to expand human rights.

The third contribution is that, instead of choosing one over another, Hasson and Ley et al. try to show how a variety of contextual and agency theories are useful for understanding neighborhood collective action. For structural explanations, they consider theories ranging from those based on universal perspectives to those based on particularistic influences. They appraise different macrosocietal factors, such as rapid urbanization and the visibility of an issue at the neighborhood level. For human agency explanations, the authors focus on kinds of neighborhood leaders, personal motivations, resources, and local political cultures. Through their

analysis they take seriously the proposition that people may respond to structure in different ways.

I suggest Hasson and Ley et al. could have improved their work in three ways. First, the authors do not explicitly define different concepts of varying importance to their analysis. One such concept is the "welfare state." As a result, Hasson and Ley et al. do not always make clear how the presence or absence of and development of the welfare state affected neighborhood group formation. For example, some of the residents of Shaughnessy Heights, an affluent neighborhood in Vancouver, endeavored to maintain the physical appearance of their neighborhood and prevent property devaluation. Was it the presence or absence of a developed Canadian welfare state that enabled residents to argue successfully that land-use regulations in Shaughnessy Heights aesthetically benefited Vancouver? Another concept Hasson and Ley et al. may have clarified, but which is less integral to their work, is "citizenship." Often considered important to analysis of welfare state development, the authors refer to citizenship and citizenship rights in the first and last chapters, but they do not thoroughly develop the notion in their eight case studies. Showing clearly how citizenship matters to neighborhood organization and local government would benefit several areas of social science research.

Second, as mentioned above, Hasson and Ley et al. do, for the most part, consider separately several theories of collective action and social change. One problem arising from this broad undertaking is that the authors cannot compare, for example, all human agency theories when examining a single neighborhood group and its actions. In a perfect world I would value their thoughts on all structural and agency theories for the four neighborhood types in each of the eight case studies.

Third, except for chapters in which neighborhoods of the same type are compared, the authors do not fully elaborate on relationships at similar levels of analysis. While examining the case studies I wondered whether, within the cities, the neighborhood organizations are aware of each other. For the most part the authors do not discuss the impact neighborhood groups have on each other.



## Book Notes

*Social Theory and the Politics of Identity*. Edited by Craig Calhoun. Cambridge, Mass.: Blackwell Publishers, 1994. Pp. viii + 350. \$21.95.

The collection of essays in *Social Theory and the Politics of Identity* grew out of a mini conference of the Theory Section of the American Sociological Association meetings in 1992. These essays attempt to capture the ways in which identity is being assessed in sociology in what is an entirely different manner than the much acknowledged approach in sociological social psychology. Traditional sociological social psychology often utilized identity terms and categories in a static fashion. On the other hand, the sociological theorists who have produced these essays have attempted to explore more critically the concept of identity as it is embedded in identity politics. This approach involves consideration of the more general discussions of subjectivity and agency, especially how individuals respond to their location in identity categories (the most common being race, class, gender, and sexual orientation) and how such responses cause the meanings and content of identity categories to change over time.

Each essay attempts to explore, criticize, or enhance a segment of the larger discussion on the political significance of identity in modern society. Essays such as Margaret R. Somers and Gloria D. Gibson's "Reclaiming the Epistemological 'Other': Narrative and the Social Construction of Identity," Charles Lemert's "Dark Thoughts about the Self," and Eli Zaretsky's "Identity Theory, Identity Politics: Psychoanalysis, Marxism, Post-Structuralism," also take into account the role of traditional sociological methods in both obscuring and augmenting scholarly research on the political significance of identity. This book provides a concise set of perspectives on the status of the politics of identity in contemporary theoretical sociology.

*Formal Theory in Sociology: Opportunity or Pitfall?* Edited by Jerald Hage. Albany: State University of New York Press, 1994. Pp. ix + 204. \$16.95.

The objective of *Formal Theory in Sociology: Opportunity or Pitfall?* is to consider, from a sociology of knowledge perspective, the demise of formal theory in sociology. Although not consistently defined throughout

notes are written by members of the book review editorial board.

by The University of Chicago. All rights reserved.

95/10006-0011\$01.50

Volume 100 Number 6

the contributions in this book, formal theory is generally taken to be that branch of sociological theorizing that more generally involves deductive modes of reasoning and, more specifically, the testing of hypotheses. The inconsistency in a precise definition for the concept leads certain contributors to identify more, and others less, promise for formal theory in contemporary sociology because they ground their comments within their own definitions.

The collection of essays here are not research-intensive endeavors. Instead, they are concise commentaries from some of the most attentive sociologists to the effort to reestablish formal theory in sociology. In the view of many of the contributors here, including Bernard Cohen, Jack Gibbs, and Hubert Blalock, formal theory was a topic of significant sociological interest in the 1960s and 1970s. These commentators argue that at that time there was a vigorous call for more research in that area. Currently, however, no such effort seems to exist. Therefore, Jerald Hage, the editor of this volume, posed to the contributors the question, "Why did such a widely based effort fail so quickly and decisively?" (p. 1). The responses that constitute this book are divided into three sections: "The Demise of Theory Construction," "Alternative Strategies for Developing Formal Theory," and "Implications for the Discipline." The book ends with a mandate for the salvation of formal theory in sociology.

*Evolutionary Economics and Chaos Theory: New Directions in Technology Studies.* Edited by Loet Leydesdorff and Peter Van den Besselaar. New York: St. Martin's Press, 1994. Pp. xiii + 215. \$55.00 (cloth); \$19.95 (paper).

This is a conference/workshop volume from the University of Amsterdam, with contributors from mathematics, management, and physics as well as the social sciences in six European countries, India, the United States, and Mexico. Its unifying theoretical theme is the extension of evolutionary approaches via models of nonlinear dynamic systems. This project, following Nelson and Winter, strengthens the link between general ecological models and the microfoundations of innovation under uncertainty (as investigated by Schumpeter, Simon, Cyert and March, etc.). The papers exhibit a variety of different takes on this idea, ranging from probabilistic models of stochastic processes to more radical "postmechanical" models of how sociocultural organizations construct their own environment in an essentially open-ended, synergistic way.

The collection is of considerable interest beyond its apparently narrow focus. The 14 papers address many issues relevant to the sociology of economic organizations: network effects, government policies, the cre-

ation of niches, the diffusion of information and norms, academia-industry links, and the translation of scientific ideas between organizations.

The papers are dominated by theory, modeling, and simulations rather than empirical research. Sociologists may be unfamiliar with some of the concepts used (e.g. punctuated equilibrium, self-organization, strange attractors, Pólya urn models, and Markov processes). However, these ideas, like the formal mathematical models used, are generally presented in a non-mystifying manner and are well couched in verbal explanations.

*New Social Movements: From Ideology to Identity.* Edited by Enrique Laraña, Hank Johnston, and Joseph R. Gusfield. Philadelphia: Temple University Press, 1994. Pp. vi + 368. \$49.95 (cloth); \$22.95 (paper).

As the title implies, *New Social Movements: From Ideology to Identity* is a collection of articles that attempts to critically assess New Social Movement theory. The editors state the objective of this collection to provide some "provisional" answers to the following questions (p. 9): (1) Are the new movements as new as they seem? (2) What social and cultural changes have led to the emergence of such movements? (3) Are the ideologies of the past 150 years, with their general programs of reform and revolution, no longer operative in these movements? and (4) Has the fulcrum of social movement action shifted from a concern for large-scale societal change to narrower, more self-oriented goals of claiming and realizing new individual and group identities?

The essays are grouped into three sections: "Culture and Identity in Contemporary Social Movements," "Collective Actors in New Social Movements," and "Collective Action and Identity in Changing Political Contexts." They respond to the questions mentioned earlier by offering perspectives on the role of identity in social movements, the place of ideology in social movements as ideology relates to collective identity, and the emergence of social issues that arise from ideational and structural continuity in efforts to mobilize collectively. Taken as a whole, this volume offers a sympathetic yet balanced account of the progress and the enduring problems in new social movement theory.

*The New Asian Immigration in Los Angeles and Global Restructuring.* Edited by Paul Ong, Edna Bonacich, and Lucie Cheng. Philadelphia: Temple University Press, 1994. Pp. xi + 330.

This collection of essays continues the discussion raised in *Labor Immigration under Capitalism: Asian Workers in the United States before*

*World War II* (edited by Lucie Cheng and Edna Bonacich [Berkeley and Los Angeles: University of California Press, 1984]) by focusing on how the post-World War II immigration of Asian-Americans to the United States has changed in light of global economic restructuring. The editors argue that the growth in number as well as diversity of Asian immigrants can be best understood in the context of macro-level changes in the Pacific Rim region. To do this, the contributors analyze the political economy of capitalist restructuring, immigration patterns, economic incorporation of Asian immigrants, and political struggles, both within the Asian population and between Asian and other racial groups in Los Angeles. Los Angeles, the editors argue, is a prime candidate for investigation because of its status as a global city, its position in the Pacific Rim, and its popularity as a city of entry for new Asian immigrants. Contributors include the editors and Steve Gold, Yen Espiritu, John Liu, and Leland Saito, among others. *The New Asian Immigration in Los Angeles and Global Restructuring* is a useful addition to the growing literature on Asian-Americans, especially in its attempt to contextualize immigration patterns within the global economy, and should be interesting to students and researchers of both Asia and Asian-Americans.

*Coping with Trouble: How Science Reacts to Political Disturbances of Research Conditions*. Edited by Uwe Schimank and Andreas Stucke. New York: St. Martin's Press, 1994. Pp. 400. \$59.95.

This volume is based on an international conference held at the Max-Planck Institute in Cologne, and contains ten case studies of how research communities react to outside threats from the political arena. The studies touch upon both university and nonuniversity research, multiple levels of interaction (including networks), funding agencies, elite bargaining, public distrust, and support from industry. They all focus on Western industrialized nations during the 1980s and 1990s, which makes them interesting from a historical and regional perspective as well as with respect to organizations and the sociology of science in general.

The editors propose a number of limited generalizations in the spirit of Merton's "grounded theory of the middle range." They formalize the concepts of "trouble" and "coping," and lay out the various analytic levels and kinds of actors involved in any given case. The approach is action-theoretic, emphasizing the complexity of interdependent reactions, and the editors suggest that generalizations can only be made at the level of constellations characteristic of particular sets of cases. For these cases, the typical pattern is one of only partially successful coping, based on the opportunistic efforts of individuals faced with the reality that research

communities have a low potential for solidaristic action, and have very little weight in the political arena.

*Sexuality across the Life Course.* Edited by Alice S. Rossi. Chicago: University of Chicago Press, 1994. Pp. xvii + 418. \$34.95.

This volume approaches sexual functioning from social behavioral, psychological, and biomedical perspectives. Recognizing sexuality as an area of human life where the physical body and the social self are inherently connected to a great degree, Alice Rossi stresses the need for the multidisciplinary approach of these papers. The work was supported by the MacArthur Foundation's Research Network on Successful Midlife Development. In seeking an age-comparative framework, however, the decision was made to investigate the full age range of sexual life from adolescence to old age. The result is a wide-ranging set of papers, organized into four substantive sections. The first section consists of only one chapter, by Rossi. This chapter draws upon evolutionary biology, behavior genetics, psychological studies of physical attractiveness, and studies of gender differences in personality in modeling both sexuality and reproduction. The second section describes the diversity of sexual behavior as it varies with history, culture, opportunity, and sexual orientation. The third section focuses on selected phases of the life course: sexual initiation in adolescence, middle age, and old age. The final section deals with health-related issues in sexuality, including the effects of chronic disease and medication on sexual functioning, as well as the nature of sexual problems and the variety of therapies used to deal with them over the past quarter-century. The chapter on the effects of chronic disease and medication illustrates the value of bringing biomedical findings to the attention of social scientists; these findings should inform researchers attempting to explain age-related declines in sexual activity. This volume is important reading for anyone involved in research on sexuality.

*Gays and Lesbians in the Military: Issues, Concerns, and Contrasts.* Edited by Wilbur J. Scott and Sandra Carson Stanley. New York: Aldine de Gruyter, 1994. Pp. xx + 278. \$43.95 (cloth); \$21.95 (paper).

The editors of *Gays and Lesbians in the Military* seek to provide a scholarly presentation of the debates surrounding the participation of homosexuals in the military. The volume provides 17 chapters written by 23 authors, all but three of whom are social scientists and more than half of whom have military experience.



Part 1 establishes the context with articles providing overviews of the political debate, gay and lesbian liberation movements, and the state of social science research on the issue. Part 2 treats the controversy as a clash of worldviews, with articles exemplifying each, plus a lead article by Laura Miller presenting original data on the diversity of soldiers' views. Part 3 assesses the analogies between the incorporation of homosexuals into the military and the previous incorporations of racial minorities and women. Part 4 provides comparisons to policy and practice in other nations (Canada, Israel, Great Britain, and the Netherlands). The last section consists of four chapters on the policy implications of lifting the ban. The collection is diverse, yet well integrated and organized. It should interest specialists in lesbian and gay studies, sociology of the military, and more broadly, status politics and social change.

AJS

*American Journal of Sociology*

Volume 100 Bimonthly July 1994–May 1995

MARTA TIENDA, Editor

JIM DAVIS, FRED KNISS, SUSAN E. MAYER, JEFFREY MORENOFF, EMILIO PARRADO, and ROBERT J. SAMPSON, Associate Editors

MARY C. BRINTON, Associate Editor, Book Review

GRACE KAO, Assistant Book Review Editor

CHAN-UNG PARK, STEPHEN B. PLANK, MICHAEL J. REAY, DAVID A. SMILDE, and ALFORD A. YOUNG, JR., Book Review Editorial Board

SUSAN ALLAN, Managing Editor

LAUREN FORD-TAGGART and CLAIRE JAMES, Editorial Assistants

**Consulting Editors** PAUL D. ALLISON • EDWIN AMENTA • PETER BEARMAN • ROBERT J. BURSİK, JR. • JOHN CASTERLINE • NANCY DAVIS • NANCY DENTON • CLAUDE FISCHER • ROBERTO FRANZOSI • ADAM GAMORAN • JOHN R. HALL • MICHAEL T. HANNAN • MICHAEL HECHTER • DOUGLAS HECKATHORN • BARBARA HEYNS • CHARLES HIRSCHMAN • K. JILL KIECOLT • TIMUR KURAN • GARY LAFREE • DAVID LAITIN • LINDA D. MOLM • DANIEL MONTI • JEYLAN T. MORTIMER • DANIEL OLSON • DUDLEY POSTON • CECILIA RIDGEWAY • CATHERINE ROSS • RUBÉN G. RUMBAUT • JOACHIM J. SAVELSBERG • DARREN E. SHERKAT • SHELLY SMITH • C. MATTHEW SNIPP • CHARLOTTE STEEH • GILLIAN STEVENS • ROBIN STRYKER • EDWARD TELLES • YU XIE

**International Consulting Editors** JOHN FOX (Canada) • MATTHIJS KALMIJN (The Netherlands) • NOAH LEWIN-EPSTEIN (Israel) • STEVEN L. RYTINA (Canada) • YUVAL YONAY (Israel)

**Editorial Board** ANDREW D. ABBOTT • GARY S. BECKER • CHARLES E. BIDWELL • DONALD J. BOGUE • RONALD BURT • TERRY N. CLARK • JAMES S. COLEMAN • JOHN L. COMAROFF • ROGER V. GOULD • WENDY GRISWOLD • EVELYN KITAGAWA • EDWARD O. LAUMANN • DONALD N. LEVINE • ALBERTO PALLONI • WILLIAM L. PARISH • MOISHE POSTONE • MARTIN RIESEBRODT • GEORGE P. STEINMETZ • ROSS M. STOLZENBERG • FRED L. STRODT-BECK • GERALD D. SUTTLES • LINDA J. WAITE • WILLIAM JULIUS WILSON • KAZUO YAMAGUCHI

THE UNIVERSITY OF CHICAGO PRESS  
CHICAGO, ILLINOIS

© 1994, 1995, by The University of Chicago

All rights reserved

Published July, September, November 1994

January, March, May 1995

CHICAGO, ILLINOIS, U.S.A.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

## ACKNOWLEDGMENTS TO REFEREES

The editors are indebted to their many colleagues who assist them in the assessment of manuscripts. The work of these referees cannot be overestimated; it is a continuous service to the profession. Following is a list of people who read at least one manuscript in 1994. We take this opportunity at the end of volume 100 to thank them.

JULIA P. ADAMS	PETER M. BLAU	THOMAS COOK
RICHARD D. ALBA	FRED L. BLOCK	PHILIP COOK
STAN L. ALBRECHT	HANS-PETER BLOSS-	WILLIAM A. CORSARO
VICTORIA D. ALEX-	FELD	JONATHAN CRANE
ANDER	KENNETH BOLLEN	MIHALY CSIKSZENTMI-
ROBERT ALTHAUSER	ROSLYN BOLOGH	HALYI
RONALD T. AMINZADE	PHILLIP BONACICH	FRANCIS T. CULLEN
NANCY AMMERMAN	ALAN BOOTH	BETH DAPONTE
LEON ANDERSON	CHRISTINE BOSE	KENNETH DAUBER
JACQUELINE ANGEL	TERRY BOSWELL	CHANDLER DAVIDSON
HELMUT K. ANHEIER	RONALD BREIGER	DEBORAH DAVIS
ROBERT J. ANTONIO	CHESTER L. BRITT III	JAMES A. DAVIS
SAID A. ARJOMAND	ROGERS BRUBAKER	GLENN DEANE
KAREN E. ASCHAF-	JOSEF BRUEDERL	MARY JO DEEGAN
FENBURG	ANTHONY BRYK	N. J. DEMERATH III
WILLIAM G. AXINN	CYNTHIA BUCKLEY	NORMAN K. DENZIN
JERALD BACHMAN	JAMES BURK	TERI DEVINE
WILLIAM BAINBRIDGE	GENE BURNS	SUSAN DE VOS
JAMES W. BALKWELL	RONALD S. BURT	ANDREAS DIEKMANN
WILLIAM P. BARNETT	JOHN SIBLEY BUTLER	PAUL DIMAGGIO
JAMES N. BARON	GLENN R. CARROLL	FRANK DOBBIN
FRANK D. BEAN	BRUCE CARRUTHERS	SANFORD DORNBUSCH
E. M. BECK	JANET SALTZMAN	MARY DOUGLAS
ULRICH BECK	CHAFETZ	ELAINE ALMA DRAPER
BETSY JANE BECKER	MARK CHAVES	ROBERT DREEBEN
HOWARD S. BECKER	LEO CHAVEZ	STEVEN C. DUBIN
PENNY A. EDGELL	ANDREW J. CHERLIN	ELLEN C. DU BOIS
BECKER	DANIEL CHIROT	THOMAS DURKIN
PIERS BEIRNE	HARVEY M. CHOLDIN	LAUREN B. EDELMAN
ROBERT N. BELLAH	MARIN E. CLARKBERG	MITCHELL L. EGGERS
MARGARET BENDROTH	ELISABETH CLEMENS	DANIEL ELAZAR
URI BEN-ELIEZER	CLIFFORD C. CLOGG	SCOTT R. ELIASON
ROBERT D. BENFORD	SUSAN H. COCHRANE	STEPHEN J. ELLINGSON
NACHMAN BEN-	LAWRENCE COHEN	CHRISTOPHER G. EL-
YEHUDA	STEVEN M. COHEN	LISON
MABEL BEREZIN	STEPHEN COLE	MUSTAFA EMIRBAYER
ALBERT J. BERGESEN	LERITA COLEMAN	BARBARA ENTWISLE
JOEL BEST	H. M. COLLINS	CYNTHIA FUCHS EP-
SUZANNE M. BIANCHI	RANDALL COLLINS	STEIN
DENISE BIELBY	JOSEF M. COLOMER	KAI T. ERIKSON
WILLIAM T. BIELBY	RAND D. CONGER	IVAN ERMAKOFF
AMY J. BINDER	ROBERT WILLIAM	KRISTIN NEUMAN ES-
MARY BLAIR-LOY	CONNELL	PINOSA
JUDITH R. BLAU	KAREN COOK	JEAN-LOUIS FABIANT

# American Journal of Sociology

- GEORGE FARKAS  
 REYNOLDS FARLEY  
 ROBERT R. FAULKNER  
 KATHERINE FAUST  
 ADRIAN FAVELL  
 JOE R. FEAGIN  
 JAMES FEARON  
 MIKE FEATHERSTONE  
 SCOTT FELD  
 PRISCILLA P. FERGUSON  
 ANNE FIGERT  
 ROGER FINKE  
 AMNON FINKELSTEIN  
 GLENN FIREBAUGH  
 NEIL FLIGSTEIN  
 CHARLES J. FOMBRUN  
 MARK A. FOSSETT  
 RANDAL S. FRANZ  
 JOHN H. FREEMAN  
 ELIOT FREIDSON  
 NOAH E. FRIEDKIN  
 DAVID PATRICK FRISBY  
 STEPHAN FUCHS  
 MARY FULBROOK  
 JOSEPH GALASKIEWICZ  
 GEORGE C. GALSTER  
 DIEGO GAMBETTA  
 WILLIAM A. GAMSON  
 HARRY B.G. GANZEBOOM  
 MADELAINE GERBAULET-VANASSE  
 ARLINE T. GERONIMUS  
 NAOMI GERSTEL  
 JANET Z. GIELE  
 THOMAS GIERYN  
 BERNHARD GIESEN  
 PEGGY C. GIORDANO  
 JENNIFER GLASS  
 NORVAL D. GLENN  
 THOMAS B. GOLD  
 MIRIAM A. GOLDEN  
 JEFFREY GOLDFARB  
 HARVEY GOLDMAN  
 FRANCES GOLDSCHIEDER  
 JACK GOLDSTONE  
 JOHN H. GOLDTHORPE  
 ERICH GOODE  
 JEFF GOODWIN
- LINDA GORDON  
 PHILIP S. GORSKI  
 MARK GOTTDIENER  
 MICHAEL R. GOTTFREDSON  
 FRANKLIN GOZA  
 MARK GRANOVETTER  
 HAROLD G. GRASMICK  
 SHERRI GRASMUCK  
 ANDREW M. GREELEY  
 DAVID F. GREENBERG  
 MARGARET E. GREENE  
 JAMES R. GREENLEY  
 LARRY J. GRIFFIN  
 AVERY M. GUEST  
 DOUGLAS T. GURAK  
 JOSEPH GUSFIELD  
 J. RICHARD HACKMAN  
 JOHN HAGAN  
 VALERIE A. HAINES  
 CHARLES N. HALABY  
 ARCHIBALD O. HALLER  
 TERENCE C. HALLIDAY  
 ULF HANNERZ  
 STEPHEN HANSELL  
 J. MARK HANSEN  
 THOMAS HANSON  
 LINGXIN HAO  
 RUSSELL HARDIN  
 SANDRA HARDING  
 ROSA HARITOS  
 DAVID L. HARVEY  
 PETER HEDSTRÖM  
 KAREN A. HEGTVEDT  
 KAREN HEIMER  
 JOHN HERITAGE  
 JOSEPH C. HERMANOWICZ  
 BRADLEY HERTTEL  
 ALEXANDER HICKS  
 DAN HILL  
 JEROME J. HIRMELSTEIN  
 ERIC L. HIRSCH  
 PAUL M. HIRSCH  
 RANDY HODSON  
 THOMAS HOFFER  
 GREGORY HOOKS  
 PATRICK M. HORAN  
 ANNE M. HORNSBY  
 DONALD L. HOROWITZ
- JAMES S. HOUSE  
 MICHAEL HOUT  
 SHERI HSEUH  
 MICHAEL HUGHES  
 SCOTT A. HUNT  
 ELLEN I. IDLER  
 ALEX INKELES  
 HIROSHI ISHIDA  
 MARY JACKMAN  
 DAVID JACOBS  
 JERRY A. JACOBS  
 PAUL JARGOWSKY  
 JAMES M. JASPER  
 GARY D. JAWORSKI  
 J. CRAIG JENKINS  
 GARY F. JENSEN  
 LEIF JENSEN  
 HANS JOAS  
 NAN E. JOHNSON  
 ROBERT ALUN JONES  
 KARA JOYNER  
 STEPHEN KALBERG  
 ARNE L. KALLEBERG  
 SHAWN KANAIAUPUNI  
 GRACE KAO  
 RASMA KARKLINS  
 LYNN A. KAROLY  
 MICHAEL KATZ  
 IRA KATZNELSON  
 VERNA M. KEITH  
 MICHAEL D. KENNEDY  
 ALAN C. KERCKHOFF  
 DAVID KETTLER  
 BARBARA STANEK KILBOURNE  
 HERBERT KITSCHOLT  
 BERT KLANDERMANS  
 REBECCA KLATCHE  
 JAMES R. KLUEGEL  
 FRED KNISS  
 DAVID KNOKE  
 KARIN D. KNORR-CETINA  
 MELVIN KOHN  
 HAGEN KOO  
 SANDERS D. KORNENMAN  
 ECE KUMBASAR  
 RICHARD W. LACHMANN  
 ROBERT LALONDE

## Acknowledgments

MICHÈLE LAMONT	J. MILLER MCPHERSON	WALTER W. POWELL
KENNETH C. LAND	ELIZABETH MENAGHAN	HARLAND PRECHEL
BRUNO LATOUR	JAMES MESSERSCHMIDT	SAMUEL PRESTON
JOHN LAUB	STEVEN F. MESSNER	THOMAS PULLUM
BRETT LAURSEN	JOHN W. MEYER	JILL QUADAGNO
YOK-SHIU F. LEE	STUART MICHAELS	ENRICO QUARANTELLI
JAEHYUCK LEE	RUTH MILKMAN	ADRIAN RAFTERY
ROBERT E. LERNER	JOHN MIROWSKY	CHARLES RAGIN
DONALD N. LEVINE	MARK S. MIZRUCHI	REBECA B. RAIJMAN
NOAH LEWIN-EPSTEIN	MANSOOR MOADDEL	JAMES RANGER-MOORE
ZAI LIANG	SUZANNE MODEL	WERNER RAUB
DANIEL T. LICHTER	JOHN MODELL	THOMAS G. RAWSKI
PAUL LICHTERMAN	ANDRE MODIGLIANI	MICHAEL J. REAY
FERNANDO LIMONGI	PHYLLIS MOEN	KENT REDDING
NAN LIN	HARVEY MOLOTCH	BARBARA RESKIN
BRUCE LINCOLN	JAMES MONTGOMERY	ERIC B. RHODES
JAMES R. LINCOLN	JEFFREY MORENOFF	MARTIN RIESEBRODT
CHARLES T. LINDHOLM	HARRIET P. MORGAN	GEORGE RITZER
BRUCE G. LINK	S. PHILIP MORGAN	ROBERT V. ROBINSON
ALLEN E. LISKA	SANDRA MORGEN	ÁKOS RÓNA-TAS
LINDA LOBAO	CALVIN MORRILL	WADE CLARK ROOF
DAVID LOCKWOOD	CHARLES MOSKOS	JAMES E. ROSENBAUM
JOHN LOFLAND	CHARLES W. MUELLER	RACHEL A. ROSENFELD
COLIN LOFTIN	CHANDRA MUKERJI	PETER H. ROSSI
JOHN R. LOGAN	EDWARD MULLEN	ALICE ROSSI
SAMUEL R. LUCAS	ROBYN L. MUNCY	GUENTHER ROTH
DAVID LUCKENBILL	MARTHA A. MYERS	DIETER RUCHT
STEVEN LUKES	DANIEL NAGIN	LLOYD I. RUDOLPH
MARK LUPHER	KATHY NECKERMAN	SUSANNE HOEBER RU-
STANFORD M. LYMAN	VICTOR G. NEE	DOLPH
MICHAEL M. MACY	LISA J. NEIDERT	DIETRICH
THOMAS MALONEY	JOYCE M. NIELSEN	RUESCHEMEYER
ROBERT MANNING	ANTHONY OBERSCHALL	MARY RUGGIE
JANE MANSBRIDGE	JOHN OGBU	STEVEN RUGGLES
MARGARET MOONEY	DENIS O'HEARN	HOWARD SACKS
MARINI	PAMELA OLIVER	MICHAEL P. SACKS
SUSAN E. MARSHALL	SUSAN OLZAK	MARSHALL D. SAHLINS
GERALD MARWELL	KARL-DIETER OPP	ROBERT SAMPSON
DOUGLAS S. MASSEY	ANN S. ORLOFF	ROBERT B. SCHAFER
JAMES L. MASSEY	MARYJANE OSA	THOMAS J. SCHEFF
ROSS L. MATSUEDA	ROBERT NASH PARKER	KATHRYN S. SCHILLER
SUSAN E. MAYER	EMILIO PARRADO	ROBERT SCHOEN
ALLAN MAZUR	ELIZA K. PAVALKO	MARK D. SCHOENHALS
DOUGLAS J. MCADAM	ANNE R. PEBLEY	L. PHILIP SCHUMM
PATRICIA L. MCCALL	BERNICE A. PESCO-	MELVIN SEEMAN
JOHN MCCARTHY	SOLIDO	STEVEN SEIDMAN
RICHARD MCCLEARY	TROND PETERSEN	RICHARD SERPE
JOAN MCCORD	RICHARD A. PETERSON	DMITRI N. SHALIN
DAVID MCFARLAND	JAMES F. PETRAS	MICHAEL J. SHANAHAN
JANE D. MCLEOD	ANDREW PICKERING	MARK SHANLEY
SCOTT GRANT McNALL	JOEL M. PODOLNY	YOSSI SHAVIT

American Journal of Sociology

- |                    |                     |                    |
|--------------------|---------------------|--------------------|
| BETH ANNE SHELTON  | JOHN R. SUTTON      | HUA-YIN WANG       |
| GI-WOOK SHIN       | C. GRAY SWICEGOOD   | KATHRYN B. WARD    |
| WESLEY SHRUM       | ANN SWIDLER         | R. STEPHEN WARNER  |
| RICHARD SHWEDER    | IVAN SZELENYI       | MARY WATERS        |
| ROBIN W. SIMON     | SZONJA SZELENYI     | ADAM WATSON        |
| JOACHIM SINGELMANN | JOAN TALBERT        | DAVID L. WEAKLIEM  |
| JITENDRA SINGH     | TONY H. TAM         | PAMELA S. WEBSTER  |
| JOHN SKVORETZ      | SIDNEY TARROW       | JEROEN WEESIE      |
| DAVID SMILDE       | MARK TAUSIG         | CLAUDE WELCH       |
| MICHAEL R. SMITH   | JAY D. TRACHMAN     | BRUCE WESTERN      |
| SHELLEY A. SMITH   | PEGGY THOITS        | HARRISON C. WHITE  |
| DAVID A. SNOW      | GEORGE THOMAS       | MICHAEL J. WHITE   |
| MICHAEL E. SOBEL   | ARLAND THORNTON     | MARTIN K. WHYTE    |
| MIN-WOONG SOHN     | MARTA TIENDA        | BRUCE WIEGAND      |
| NADER SOHRABI      | CHARLES TILLY       | HAROLD L. WILENSKY |
| KAROL SOLTAN       | STEVEN M. TIPTON    | DAVID R. WILLIAMS  |
| AAGE SØRENSEN      | EDWARD A. TIRYAKIAN | RHYS H. WILLIAMS   |
| ANNEMETTE SØRENSEN | CHARLES R. TITTLE   | CHARLES V. WILLIE  |
| SEYMOUR SPILERMAN  | CHARLES M. TOLBERT  | JOHN R. WILMOTH    |
| LYNETTE SPILLMAN   | III                 | DIANE WOLF         |
| GLENNA SPITZE      | JONATHAN H. TURNER  | ALAN WOLFE         |
| SUZANNE STAG-      | RALPH H. TURNER     | KURT H. WOLFF      |
| GENBORG            | STEPHEN TURNER      | RAYMOND SIN-KWOK   |
| SUSAN LEIGH STAR   | J. RICHARD UDRY     | WONG               |
| LINDA BREWSTER     | EDWINA UEHARA       | ERIK OLIN WRIGHT   |
| STEARNS            | DEBRA UMBERSON      | LAWRENCE L. WU     |
| COOKIE W. STEPHAN  | WITCHŪ VEJJAJIVA    | ROBERT WUTHNOW     |
| HAYA STIER         | LISE VOGEL          | TOSHIO YAMAGISHI   |
| STEPHEN M. STIGLER | THOMAS VOSS         | KAZUO YAMAGUCHI    |
| ARTHUR STINCHCOMBE | LINDA J. WAITE      | DALI YANG          |
| CLARENCE STONE     | ANDREW G. WALDER    | MAYER N. ZALD      |
| DAVID STRANG       | ROGER WALDINGER     | EVIATAR ZERUBAVEL  |
| SHELDON STRYKER    | HENRY A. WALKER     | EKKART ZIMMERMANN  |
| GENE F. SUMMERS    | MICHAEL WALLACE     | VERA ZOLBERG       |

## CONTENTS OF VOLUME 100

### Articles

- 1294 ALLISON, PAUL D. The Impact of Random Predictors on Comparisons of Coefficients between Models: Comment on Clogg, Petkova, and Haritou
- 859 ANHEIER, HELMUT K., JÜRGEN GERHARDS, and FRANK P. ROMO. Forms of Capital and Social Structure in Cultural Fields: Examining Bourdieu's Social Topography
- 381 BARRON, DAVID N., ELIZABETH WEST, and MICHAEL T. HANNAN. A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York City, 1914–1990
- 477 BATCHELDER, WILLIAM H. See KUMBASAR
- 346 BAUM, JOEL A. C., and JITENDRA V. SINGH. Organizational Niches and the Dynamics of Organizational Mortality
- 689 BERON, KURT. See KILBOURNE
- 781 BLAU, JUDITH R. See LAND
- 1028 BOEHNKE, KLAUS. See HAGAN
- 1479 BOSWELL, TERRY. See BROWN
- 652 BRINES, JULIE. Economic Dependency, Gender, and the Division of Labor at Home
- 1099 BRINTON, MARY C., YEAN-JU LEE, and WILLIAM L. PARISH. Married Women's Employment in Rapidly Industrializing Societies: Examples from East Asia
- 1479 BROWN, CLIFF, and TERRY BOSWELL. Strikebreaking or Solidarity in the Great Steel Strike of 1919: A Split Labor Market, Game-Theoretic, and QCA Analysis
- 720 CARROLL, GLENN R., and J. RICHARD HARRISON. On the Historical Efficiency of Competition between Organizational Populations
- 931 CHARLES, MARIA, and DAVID B. GRUSKY. Models for Describing the Underlying Structure of Sex Segregation
- 972 CHEN, KEVIN. See FIREBAUGH
- 1261 CLOGG, CLIFFORD C., EVA PETKOVA, and ADAMANTIOS HARITOU. Statistical Methods for Comparing Regression Coefficients between Models



- 1305 CLOGG, CLIFFORD C., and EVA PETKOVA, with the assistance of TZUWEI CHENG. Reply to Allison: More on Comparing Regression Coefficients
- 104 COHEN, LISA E. See HAVEMAN
- 1616 COLEMAN, JAMES S. Comment on Kuran and Collins
- 1552 COLLINS, RANDALL. Prediction in Macrosociology: The Case of the Soviet Collapse
- 1448 DAVIES, SCOTT. Leaps of Faith: Shifting Currents in Critical Sociology of Education
- 781 DAVIS, WALTER R. See LAND
- 997 DE GRAAF, NAN DIRK, PAUL NIEUWBEERTA, and ANTHONY HEATH. Class Mobility and Political Preferences: Individual and Contextual Effects
- 1157 DIXON, JO. The Organizational Context of Criminal Sentencing
- 689 ENGLAND, PAULA. See KILBOURNE
- 689 FARKAS, GEORGE. See KILBOURNE
- 750 FARLEY, REYNOLDS, CHARLOTTE STEEH, MARIA KRYSAN, TARA JACKSON, and KEITH REEVES. Stereotypes and Segregation: Neighborhoods in the Detroit Area
- 972 FIREBAUGH, GLENN, and KEVIN CHEN. Vote Turnout of Nineteenth Amendment Women: The Enduring Effect of Disenfranchisement
- 859 GERHARDS, JÜRGEN. See ANHEIER
- 931 GRUSKY, DAVID B. See CHARLES
- 1028 HAGAN, JOHN, HANS MERKENS, and KLAUS BOEHNKE. Delinquency and Disdain: Social Capital and the Control of Right-Wing Extremism among East and West Berlin Youth
- 381 HANNAN, MICHAEL T. See BARRON
- 1261 HARITOU, ADAMANTIOS. See CLOGG
- 720 HARRISON, J. RICHARD. See CARROLL
- 104 HAVEMAN, HEATHER A., and LISA E. COHEN. The Ecological Dynamics of Careers: The Impact of Organizational Founding, Dissolution, and Merger on Job Mobility
- 70 HAYNES, STEPHEN E., and DAVID JACOBS. Macroeconomics, Economic Stratification, and Partisanship: A Longitudinal Analysis of Contingent Shifts in Political Identification
- 997 HEATH, ANTHONY. See DE GRAAF
- 1682

- 1520 HECHTER, MICHAEL. Reflections on Historical Prophecy in the Social Sciences
- 750 JACKSON, TARA. See FARLEY
- 70 JACOBS, DAVID. See HAYNES
- 1 JONES, ROBERT ALUN. Ambivalent Cartesians: Durkheim, Montesquieu, and Method
- 422 KALMIJN, MATTHIJS. Assortative Mating by Cultural and Economic Occupational Status
- 689 KILBOURNE, BARBARA STANEK, PAULA ENGLAND, GEORGE FARKAS, KURT BERON, and DOROTHEA WEIR. Returns to Skill, Compensating Differentials, and Gender Bias: Effects of Occupational Characteristics on the Wages of White Women and Men
- 1611 KISER, EDGAR. What Can Sociological Theories Predict? Comment on Collins, Kuran, and Tilly
- 313 KOLLOCK, PETER. The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust
- 750 KRYSAN, MARIA. See FARLEY
- 477 KUMBASAR, ECE, A. KIMBALL ROMNEY, and WILLIAM H. BATCHELDER. Systematic Biases in Social Perception
- 1528 KURAN, TIMUR. The Inevitability of Future Revolutionary Surprises
- 781 LAND, KENNETH C., WALTER R. DAVIS, and JUDITH R. BLAU. Organizing the Boys of Summer: The Evolution of U.S. Minor-League Baseball, 1883–1990
- 1099 LEE, YEAN-JU. See BRINTON
- 904 LEHMANN, JENNIFER M. Durkheim's Theories of Deviance and Suicide: A Feminist Reconsideration
- 506 LENNON, MARY CLARE, and SARAH ROSENFELD. Relative Fairness and the Division of Housework: The Importance of Options
- 622 LEWIN-EPSTEIN, NOAH, and MOSHE SEMYONOV. Sheltered Labor Markets, Public Sector Employment, and Socioeconomic Returns to Education of Arabs in Israel
- 1131 LILLARD, LEE A., and LINDA J. WAITE. 'Til Death Do Us Part: Marital Disruption and Mortality
- 453 LOGAN, JOHN R., and GLENNA D. SPITZE. Family Neighbors
- 1028 MERKENS, HANS. See HAGAN

American Journal of Sociology

- 997 NIEUWBEERTA, PAUL. See DE GRAAF
- 587 O'HEARN, DENIS. Innovation and the World-System Hierarchy: British Subjugation of the Irish Cotton Industry, 1780–1830
- 196 OLZAK, SUSAN, SUZANNE SHANAHAN, and ELIZABETH WEST. School Desegregation, Interracial Exposure, and Antibusing Activity in Contemporary Urban America
- 153 PAMPEL, FRED C. Population Aging, Class Context, and Age Inequality in Public Spending
- 1099 PARISH, WILLIAM L. See BRINTON
- 1261 PETKOVA, EVA. See CLOGG
- 1305 PETKOVA, EVA. See CLOGG
- 1224 PODOLNY, JOEL M., and TOBY E. STUART. A Role-Based Ecology of Technological Change
- 1620 PORTES, ALEJANDRO. On Grand Surprises and Modest Certainties: Comment on Kuran, Collins, and Tilly
- 750 REEVES, KEITH. See FARLEY
- 477 ROMNEY, A. KIMBALL. See KUMBASAR
- 859 ROMO, FRANK P. See ANHEIER
- 40 RÓNA-TAS, ÁKOS. The First Shall Be Last? Entrepreneurship and Communist Cadres in the Transition from Socialism
- 506 ROSENFELD, SARAH. See LENNON
- 622 SEMYONOV, MOSHE. See LEWIN-EPSTEIN
- 196 SHANAHAN, SUZANNE. See OLZAK
- 346 SINGH, JITENDRA V. See BAUM
- ix SMALL, ALBION. The Era of Sociology
- 1383 SOHRABI, NADER. Historicizing Revolutions: Constitutional Revolutions in the Ottoman Empire, Iran, and Russia, 1905–1908
- 453 SPITZE, GLENNA D. See LOGAN
- 750 STEEH, CHARLOTTE. See FARLEY
- 1224 STUART, TOBY E. See PODOLNY
- 1199 TELLES, EDWARD E. Structural Sources of Socioeconomic Segregation in Brazilian Metropolitan Areas
- 1594 TILLY, CHARLES. To Explain Political Processes
- 1684

- 1131 WAITE, LINDA J. See LILLARD
- 689 WEIR, DOROTHEA. See KILBOURNE
- 196 WEST, ELIZABETH. See OLZAK
- 381 WEST, ELIZABETH. See BARRON

### Commentary and Debate

- 1313 BESNARD, PHILIPPE. The Study of Social Taste through First Names: Comment on Lieberman and Bell
- 1317 LIEBERSON, STANLEY. Response to Philippe Besnard
- 1333 MASSEY, DOUGLAS S., and MITCHELL L. EGGERS. Beyond the Technical Details: Reply to St. John
- 1325 ST. JOHN, CRAIG. Interclass Segregation, Poverty, and Poverty Concentration: Comment on Massey and Eggers

### Book Reviews

- 835 ABRAMS, DOUGLAS M. *Conflict, Competition, or Cooperation? Dilemmas of State Education Policymaking*. Peter W. Cookson, Jr.
- 556 AGGER, BEN. *Gender, Culture, and Power: Toward a Feminist Postmodern Critical Theory*. Charles C. Lemert
- 273 AHMAD, AIJAZ. *In Theory: Classes, Nations, and Literatures*. S. M. Shamsul Alam
- 247 AKENSON, DONALD HARMAN. *God's Peoples: Covenant and Land in South Africa, Israel, and Ulster*. José Casanova
- 1071 ALLISON, ANNE. *Nightwork: Sexuality, Pleasure, and Corporate Masculinity in a Tokyo Hostess Club*. Yuko Ogasawara
- 275 AMINZADE, RONALD. *Ballots and Barricades: Class Formation and Republican Politics in France, 1830-1871*. Samuel Cohn
- 1644 ANDERSON, LESLIE E. *The Political Ecology of the Modern Peasant: Calculation and Community*. Marc Edelman
- 841 ANGEL, RONALD J., and JACQUELINE L. ANGEL. *Painful Inheritance: Health and the New Generation of Fatherless Families*. John R. Weeks
- 1082 ANSPACH, RENEE R. *Deciding Who Lives: Fateful Choices in the Intensive-Care Nursery*. Charles Bosk

- 1346 APPELBAUM, EILEEN, and ROSEMARY BATT. *The New American Workplace: Transforming Work Systems in the United States*. Vicki Smith
- 305 BACHMAN, RONET. *Death and Violence on the Reservation: Homicide, Family Violence, and Suicide in American Indian Populations*. Philip A. May
- 292 BARNETT, LARRY D. *Legal Construct, Social Concept: A Macrosociological Perspective on Law*. Robin Stryker
- 533 BEARMAN, PETER S. *Relations into Rhetorics: Local Elite Social Structure in Norfolk, England, 1540-1640*. Christopher K. Ansell
- 1347 BECK, JOHN C., and MARTHA N. BECK. *The Change of a Lifetime: Employment Patterns among Japan's Managerial Elite*. Hiroshi Ishida
- 1648 BELL, MICHAEL MAYERFELD. *Childerley: Nature and Morality in a Country Village*. Wendy Griswold
- 585 BELLER, ANDREA H., and JOHN W. GRAHAM. *Small Change: The Economics of Child Support*. Jay D. Teachman
- 561 BERGMANN, JÖRG R. *Discreet Indiscretions: The Social Organization of Gossip*. Stephen Ellingson
- 828 BIAN, YANJIE. *Work and Inequality in Urban China*. Tak Chuen Luk
- 1634 BLAIKIE, ANDREW. *Illegitimacy, Sex and Society: Northeast Scotland, 1750-1900*. Peter Becker
- 563 BLAU, JUDITH R. *Social Contracts and Economic Markets*. Richard Swedberg
- 309 BORDO, SUSAN. *Unbearable Weight: Feminism, Western Culture, and the Body*. Sue Fisher
- 1079 BRODIE, JANET FARRELL. *Contraception and Abortion in Nineteenth-Century America*. Carl N. Degler
- 829 BUGRA, AYSE. *State and Business in Modern Turkey: A Comparative Study*. Roderic Ai Camp
- 1062 BURDICK, JOHN. *Looking for God in Brazil: The Progressive Catholic Church in Urban Brazil's Religious Arena*. David A. Smilde
- 265 CHALOUPEK, WILLIAM. *Knowing Nukes: The Politics and Culture of the Atom*. Andrew Marchant-Shapiro
- 1636 CHAUNCEY, GEORGE. *Gay New York: Gender, Urban Culture, and the Making of the Gay Male World, 1890-1940*. Vern L. Bullough

- 1073 CHODOROW, NANCY J. *Femininities, Masculinities, Sexualities: Freud and Beyond*. Patricia Ticineto Clough
- 535 COHN, SAMUEL. *When Strikes Make Sense—and Why: Lessons from Third Republic French Coal Miners*. Michael Wallace
- 1630 COOKE, MAEVE. *Language and Reason: A Study of Habermas's Pragmatics*. James J. Chriss
- 837 COOKSON, PETER W., JR. *School Choice: The Struggle for the Soul of American Education*. Stephen B. Plank
- 541 COX, ROBERT H. *The Development of the Dutch Welfare State: From Workers' Insurance to Universal Entitlement*. Joya Misra
- 551 DAVIS, FRED. *Fashion, Culture, and Identity*. Valerie Steele
- 1627 DEAN, MITCHELL. *Critical and Effective Histories: Foucault's Methods and Historical Sociology*. Barry Smart
- 1640 DENOEU, GUILAIN. *Urban Unrest in the Middle East: A Comparative Study of Informal Networks in Egypt, Iran, and Lebanon*. Nader Sohrabi
- 1058 DOBBIN, FRANK. *Forging Industrial Policy: The United States, Britain, and France in the Railway Age*. Richard Rubinson
- 1663 DOUGHERTY, KEVIN J. *The Contradictory College: The Conflict-ing Origins, Impacts, and Futures of the Community College*. Walter R. Allen
- 547 ECHEVERRI-GENT, JOHN. *The State and the Poor: Public Policy and Political Development in India and the United States*. Patrick Heller
- 249 EDER, KLAUS. *The New Politics of Class: Social Movements and Cultural Dynamics in Advanced Societies*. Hanspeter Kriesi
- 1343 EINHORN, BARBARA. *Cinderella Goes to Market: Citizenship, Gender, and Women's Movements in East Central Europe*. Barbara Heyns
- 296 ENNIS, PHILIP H. *The Seventh Stream: The Emergence of Rockn-roll in American Popular Music*. Richard A. Peterson
- 252 ERIKSON, ROBERT, and JOHN H. GOLDTHORPE. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Hartmut Kaelble
- 1091 FLANAGAN, WILLIAM G. *Contemporary Urban Sociology*. Harvey M. Choldin
- 282 FLEMMING, ROY B., PETER F. NARDULLI, and JAMES EISEN-

- STEIN. *The Craft of Justice: Politics and Work in Criminal Court Communities*. Celesta A. Albonetti
- 301 FOX, RENÉE C., and JUDITH P. SWAZEY, with JUDITH C. WATKINS. *Spare Parts: Organ Replacement in American Society*. Gail Henderson
- 831 GAMBETTA, DIEGO. *The Sicilian Mafia: The Business of Private Protection*. Peter Kollock
- 1069 GAMSON, JOSHUA. *Claims to Fame: Celebrity in Contemporary America*. Michèle Lamont
- 582 GERSON, KATHLEEN. *No Man's Land: Men's Changing Commitments to Family and Work*. Bob Blauner
- 1358 GOLDSCHIEDER, FRANCES K., and CALVIN GOLDSCHIEDER. *Leaving Home before Marriage: Ethnicity, Familism, and Generational Relationships*. William G. Axinn
- 1068 HALLE, DAVID. *Inside Culture: Art and Class in the American Home*. Eugene Halton
- 289 HALLIDAY, TERENCE C., and MORRIS JANOWITZ, eds. *Sociology and Its Publics: The Forms and Fates of Disciplinary Organization*. Howard Schuman
- 1637 HANCHARD, MICHAEL GEORGE. *Orpheus and Power: The "Movimento Negro" of Rio de Janeiro and São Paulo, Brazil, 1945–1988*. Richard Graham
- 242 HANNERZ, ULF. *Cultural Complexity: Studies in the Social Organization of Meaning*. Wendy Griswold
- 1338 HANSEN, KAREN V. *A Very Social Time: Crafting Community in Antebellum New England*. Jeanne Boydston
- 256 HART, JEFFREY A. *Rival Capitalists: International Competitiveness in the United States, Japan, and Western Europe*. Alan Cawson
- 1665 HASSON, SHLOMO, and DAVID LEY. *Neighbourhood Organizations and the Welfare State*. Brian K. Gran
- 1655 HAVEMAN, ROBERT H., and BARBARA L. WOLFE. *Succeeding Generations: On the Effects of Investments in Children*. Susan E. Mayer
- 1066 HAYS, KIM. *Practicing Virtues: Moral Traditions at Quaker and Military Boarding Schools*. Fred Kniss
- 1653 HAZAN, HAIM. *Old Age: Constructions and Deconstructions*. Anne Foner
- 1688

- 287 HEINZ, JOHN P., EDWARD O. LAUMANN, ROBERT L. NELSON, and ROBERT H. SALISBURY. *The Hollow Core: Private Interests in National Policy Making*. Paul A. Sabatier
- 253 HODGE, ROBERT W., and NAOHIRO OGAWA. *Fertility Change in Contemporary Japan*. Masanori Hashimoto
- 1056 HOLLINGSWORTH, J. ROGERS, PHILLIPE C. SCHMITTER, and WOLFGANG STREECK, eds. *Governing Capitalist Economies: Performance and Control of Economic Sectors*. Peter A. Hall
- 281 HOLTZBERG-CALL, MAGGIE. *The Lost World of the Craft Printer*. Robert Max Jackson
- 824 HOROWITZ, IRVING LOUIS. *The Decomposition of Sociology*. David G. Wagner
- 817 HUFF, TOBY E. *The Rise of Early Modern Science: Islam, China, and the West*. Benjamin A. Elman
- 1642 IRELAND, PATRICK. *The Policy Challenge of Ethnic Diversity: Immigrant Politics in France and Switzerland*. Carl-Ulrik Schierup
- 822 JACKMAN, ROBERT W. *Power without Force: The Political Capacity of Nation-States*. Michael Barnett
- 532 JANSSENS, ANGÉLIQUE. *Family and Social Change: The Household as a Process in an Industrializing Community*. Steven Ruggles
- 279 JELLISON, KATHERINE. *Entitled to Power: Farm Women and Technology, 1913-1963*. Jane M. Pederson
- 1657 JESSOR, RICHARD, JOHN EDWARD DONOVAN, and FRANCES MARIE COSTA. *Beyond Adolescence: Problem Behavior and Young Adult Development*. Robert B. Cairns
- 261 JOPPKE, CHRISTIAN. *Mobilizing against Nuclear Energy: A Comparison of Germany and the United States*. David S. Meyer
- 564 JORDAN, WILLIAM CHESTER. *Women and Credit in Pre-Industrial and Developing Societies*. Elaine Kempson
- 1646 JUDD, ELLEN R. *Gender and Power in Rural North China*. Kenneth Pomeranz
- 821 KALBERG, STEPHEN. *Max Weber's Comparative-Historical Sociology*. Keith Tribe
- 576 KERCKHOFF, ALAN C. *Diverging Pathways: Social Structure and Career Deflections*. Aage B. Sørensen
- 260 KLEIDMAN, ROBERT. *Organizing for Peace: Neutrality, the Test Ban, and the Freeze*. James M. Jasper



- 1360 KLINE, STEPHEN. *Out of the Garden: Toys and Children's Culture in the Age of TV Marketing*. Gary Alan Fine
- 826 KOO, HAGEN, ed. *State and Society in Contemporary Korea*. Gi-Wook Shin
- 553 KOSMIN, BARRY A., and SEYMOUR P. LACHMAN. *One Nation under God: Religion in Contemporary American Society*. Patricia M. Y. Chang
- 1364 KOWALEWSKI, MARK R. *All Things to All People: The Catholic Church Confronts the AIDS Crisis*. Jim Wolf
- 539 KRIESI, HANSPETER. *Political Mobilization and Social Change: The Dutch Case in Comparative Perspective*. Julia P. Adams
- 1074 KUNZEL, REGINA G. *Fallen Women, Problem Girls: Unmarried Mothers and the Professionalization of Social Work, 1890-1945*. Kristin Kay Barker
- 300 LAFOLLETTE, MARCEL C. *Stealing into Print: Fraud, Plagiarism, and Misconduct in Scientific Publishing*. Allan Mazur
- 244 LAMONT, MICHÈLE, and MARCEL FOURNIER, eds. *Cultivating Differences: Symbolic Boundaries and the Making of Inequality*. Tia Denora
- 566 LEBRA, TAKIE SUGIYAMA. *Above the Clouds: Status Culture of the Modern Japanese Nobility*. Jane M. Bachnik
- 1375 LEHMANN, JENNIFER M. *Durkheim and Women*. Edward A. Tiryakian
- 571 LEWIN-EPSTEIN, NOAH, and MOSHE SEMYONOV. *The Arab Minority in Israel's Economy: Patterns of Ethnic Inequality*. Calvin Goldscheider
- 1340 LOCK, MARGARET. *Encounters with Aging: Mythologies of Menopause in Japan and North America*. L. L. Cornell
- 1336 LORBER, JUDITH. *Paradoxes of Gender*. Rosanna Hertz
- 303 LUHMANN, NIKLAS. *Risk: A Sociological Theory*. Alfonso Morales
- 819 MANN, MICHAEL. *The Sources of Social Power*. Volume 2: *The Rise of Classes and Nation-States, 1760-1914*. Chris Sparks
- 1356 MARGOLIN, LESLIE. *Goodness Personified: The Emergence of Gifted Children*. Spencer E. Cahill
- 1651 MARTIN, LINDA G., and SAMUEL H. PRESTON, eds. *Demography of Aging*. İşik A. Aytaç
- 1690

- 307 MESSERSCHMIDT, JAMES W. *Masculinities and Crime: Critique and Reconceptualization of Theory*. Kathleen Daly
- 1365 MILLER, JON. *The Social Control of Religious Zeal: A Study of Organizational Contradictions*. Adam B. Seligman
- 294 NOCK, STEVEN L. *The Costs of Privacy: Surveillance and Reputation in America*. Richard V. Ericson
- 578 OKANO, KAORI. *School to Work Transition in Japan: An Ethnographic Study*. Mamoru Tsukada
- 545 ORLOFF, ANN SHOLA. *The Politics of Pensions: A Comparative Analysis of Britain, Canada, and the United States, 1880-1940*. Jens Alber
- 568 OXFELD, ELLEN. *Blood, Sweat, and Mahjong: Family and Enterprise in an Overseas Chinese Community*. Martin King Whyte
- 1660 PARCEL, TOBY L., and ELIZABETH G. MENAGHAN. *Parents' Jobs and Children's Lives*. Janet Currie
- 1662 PARKER, ROBERT E. *Flesh Peddlers and Warm Bodies: The Temporary Help Industry and Its Workers*. Alison Davis-Blake
- 269 PÉREZ-DÍAZ, VICTOR. *The Return of Civil Society; The Emergence of Democratic Spain*. Felipe Pimentel
- 1086 PERRUCCI, ROBERT. *Japanese Auto Transplants in the Heartland: Corporatism and Community*. Michael Schwartz
- 271 PONTUSSON, JONAS. *The Limits of Social Democracy: Investment Politics in Sweden*. Madelaine Gerbaulet-Vanasse
- 570 PORTES, ALEJANDRO, and ALEX STEPICK. *City on the Edge: The Transformation of Miami*. Thomas J. Espenshade
- 843 PRATT, HENRY J. *Gray Agendas: Interest Groups and Public Pensions in Canada, Britain, and the United States*. Hugh Heclo
- 246 RIESEBRODT, MARTIN. *Pious Passion: The Emergence of Modern Fundamentalism in the United States and Iran*. Steve Bruce
- 1341 ROPER, MICHAEL. *Masculinity and the British Organization Man since 1945*. Mary Blair-Loy
- 1093 ROSENTHAL, ROB. *Homeless in Paradise: A Map of the Terrain*. Daniel M. Cress
- 1064 RUBIN, JULIUS H. *Religious Melancholy and Protestant Experience in America*. Richard W. Flory
- 267 RUESCHEMEYER, DIETRICH, EVELYNE HUBER STEPHENS, and

- JOHN D. STEPHENS. *Capitalist Development and Democracy*. Dongno Kim
- 537 SANDOVAL, SALVADOR A. M. *Social Change and Labor Unrest in Brazil since 1945*. Gay W. Seidman
- 1094 SAVELSBERG, JOACHIM, with PETER BRUHL. *Constructing White-Collar Crime: Rationalities, Communication, Power*. James William Coleman
- 1629 SCHMAUS, WARREN. *Durkheim's Philosophy of Science and the Sociology of Knowledge: Creating an Intellectual Niche*. Mark A. Schneider
- 839 SCHNEIDER, BARBARA, and JAMES S. COLEMAN, eds. *Parents, Their Children, and Schools*. Doris R. Entwisle
- 814 SCHNEIDER, MARK A. *Culture and Enchantment*. Chandra Mukerji
- 1060 SEIDMAN, GAY W. *Manufacturing Militance: Workers' Movements in Brazil and South Africa, 1970-1985*. Kevin Neuhouser
- 258 SHALEV, MICHAEL. *Labour and the Political Economy in Israel*. Ilan Talmud
- 1374 SHAPIN, STEVEN. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Kevin L. Brown
- 573 SHAVIT, YOSHI, and HANS-PETER BLOSSFELD, eds. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Kathryn S. Schiller
- 1096 SIMON, JONATHAN. *Poor Discipline: Parole and the Social Control of the Underclass, 1890-1990*. John R. Sutton
- 277 SKOCPOL, THEDA. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. John Mohr
- 558 SMITH, THOMAS SPENCE. *Strong Interaction*. Norman K. Denzin
- 583 SNAREY, JOHN. *How Fathers Care for the Next Generation: A Four-Decade Study*. Donald J. Hernandez
- 1351 SNIDERMAN, PAUL M., and THOMAS PIAZZA. *The Scar of Race*. David O. Sears
- 833 STANLEY, ROBERT. *Dimensions of Law in the Service of Order: Origins of the Federal Income Tax, 1861-1913*. John L. Campbell
- 1055 STEINMETZ, GEORGE. *Regulating the Social: The Welfare State and Local Politics in Imperial Germany*. Abram de Swaan
- 1371 STERNHELL, ZEEV, with MARIO SZNAJDER and MAIA ASHERI.
- 1692

- The Birth of Fascist Ideology: From Cultural Rebellion to Political Revolution.* Stephen Turner
- 1089 STOECKER, RANDY. *Defending Community: The Struggle for Alternative Development in Cedar-Riverside.* Alford A. Young, Jr.
- 554 STRAUSS, ANSELM L. *Continual Permutations of Action.* Thomas Spence Smith
- 1632 SUMNER, COLIN. *The Sociology of Deviance: An Obituary.* Sudhir Alladi Venkatesh
- 1349 THELEN, KATHLEEN A. *Union of Parts: Labor Politics in Post-war Germany.* Colin Crouch
- 1084 THOMAS, ROBERT J. *What Machines Can't Do: Politics and Technology in the Industrial Enterprise.* Steven Peter Vallas
- 580 TOMASKOVIC-DEVEY, DONALD. *Gender and Racial Inequality at Work: The Sources and Consequences of Job Segregation.* Patricia A. Gwartney-Gibbs
- 1080 TRENT, JAMES W., JR. *Inventing the Feeble Mind: A History of Mental Retardation in the United States.* Colin Samson
- 1053 VOSS, KIM. *The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century.* Jeffrey Haydu
- 847 WAGNER, DAVID. *Checkerboard Square: Culture and Resistance in a Homeless Community.* J. William Spencer
- 848 WAGNER-PACIFICI, ROBIN. *Discourse and Destruction: The City of Philadelphia versus MOVE.* Marc W. Steinberg
- 845 WALSH, JOHN P. *Supermarkets Transformed: Understanding Organizational and Technological Innovations.* Peter Birkeland
- 1077 WARD, W. PETER. *Birth Weight and Economic Growth: Women's Living Standards in the Industrializing West.* Jeremy Atack
- 851 WEINBERG, MARTIN S., COLIN J. WILLIAMS, and DOUGLAS W. PRYOR. *Dual Attraction: Understanding Bisexuality.* Paula C. Rust
- 1362 WHITE, HARRISON C. *Careers and Creativity: Social Forces in the Arts.* Diana Crane
- 1369 WIGGERSHAUS, ROLF. *The Frankfurt School: Its History, Theories, and Political Significance.* Douglas Kellner
- 543 WILLIAMSON, JOHN B., and FRED C. PAMPEL. *Old-Age Security in Comparative Perspective.* Tang Nah Ng and Alexander Hicks

American Journal of Sociology

- 285 WOLMAN, HAROLD, and MICHAEL GOLDSMITH. *Urban Politics and Policy: A Comparative Approach*. Edward J. Blakely
- 1367 WRONG, DENNIS H. *The Problem of Order: What Unites and Divides Society*. Satoshi Kanazawa
- 1354 WUTHNOW, ROBERT. *Sharing the Journey: Support Groups and America's New Quest for Community*. David M. Hummon
- 549 ZELIZER, BARBIE. *Covering the Body: The Kennedy Assassination, the Media, and the Shaping of Collective Memory*. Daniel C. Hallin
- 298 ZEY, MARY. *Banking on Fraud: Drexel, Junk Bonds, and Buy-outs*. Kevin J. Delaney
- 263 ZISK, KIMBERLY MARTEN. *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991*. Roger Hamburg

Book Notes

- 1378 *After Ellis Island: Newcomers and Natives in the 1910 Census*, edited by SUSAN COTTS WATKINS
- 1380 *Authority of the Consumer*, edited by RUSSELL KEAT, NIGEL WHITELEY, and NICHOLAS ABERCROMBIE
- 856 *Bourdieu: Critical Perspectives*, edited by CRAIG CALHOUN, EDWARD LIPUMA, and MOISHE POSTONE
- 1379 *Chinese Families in the Post-Mao Era*, edited by DEBORAH DAVIS and STEVAN HARRELL
- 1671 *Coping with Trouble: How Science Reacts to Political Disturbances of Research Conditions*, edited by UWE SCHIMANK and ANDREAS STUCKE
- 1380 *Cultures in Flux: Lower-Class Values, Practices, and Resistance in Late Imperial Russia*, edited by STEPHEN P. FRANK and MARK D. STEINBERG
- 1378 *Democratic Reform and the Position of Women in Transitional Economies*, edited by VALENTINE M. MOGHADAM
- 1669 *Evolutionary Economics and Chaos Theory: New Directions in Technology Studies*, edited by LOET LEYDESDORFF and PETER VAN DEN BESSELAAR
- 1668 *Formal Theory in Sociology: Opportunity or Pitfall?* edited by JERALD HAGE

- 1672 *Gays and Lesbians in the Military: Issues, Concerns, and Contrasts*, edited by WILBUR J. SCOTT and SANDRA CARSON STANLEY
- 855 *Immigration and Ethnicity: The Integration of America's Newest Arrivals*, edited by BARRY EDMONSTON and JEFFREY S. PASSEL
- 854 *In the Barrios: Latinos and the Underclass Debate*, edited by JOAN MOORE and RAQUEL PINDERHUGHES
- 1670 *New Asian Immigration in Los Angeles and Global Restructuring*, edited by PAUL ONG, EDNA BONACICH, and LUCIE CHENG
- 854 *Newcomers in the Workplace: Immigrants and the Restructuring of the U.S. Economy*, edited by LOUISE LAMPHERE, ALEX STEPHICK, and GUILLERMO GERNIER
- 1670 *New Social Movements: From Ideology to Identity*, edited by ENRIQUE LARAÑA, HANK JOHNSTON, and JOSEPH R. GUSFIELD
- 857 *Paying the Piper: Causes and Consequences of Art Patronage*, edited by JUDITH HUGGINS BALFE
- 855 *Researching Women's Lives from a Feminist Perspective*, edited by MARY MAYNARD and JUNE PURVIS
- 1672 *Sexuality across the Life Course*, edited by ALICE S. ROSSI
- 1668 *Social Theory and the Politics of Identity*, edited by CRAIG CALHOUN
- 857 *Theoretical Research Programs: Studies in the Growth of Theory*, edited by JOSEPH BERGER and MORRIS ZELDITCH, JR.
- 1380 *Workplace Industrial Relations and the Global Challenge*, edited by JACQUES BELANGER, P. K. EDWARDS, and LARRY HAIVEN

## BOOK REVIEWERS

- JULIA P. ADAMS, 539  
S. M. SHAMSUL ALAM, 273  
JENS ALBER, 545  
CELESTA A. ALBONETTI, 282  
WALTER R. ALLEN, 1663  
CHRISTOPHER K. ANSELL, 533  
JEREMY ATACK, 1077  
WILLIAM G. AXINN, 1358  
IŞIK A. AYTAÇ, 1651  
JANE M. BACHNIK, 566  
KRISTIN KAY BARKER, 1074  
MICHAEL BARNETT, 822  
PETER BECKER, 1634  
PETER BIRKELAND, 845  
MARY BLAIR-LOY, 1341  
EDWARD J. BLAKELY, 285  
BOB BLAUNER, 582  
CHARLES BOSK, 1082  
JEANNE BOYDSTON, 1338  
KEVIN L. BROWN, 1374  
STEVE BRUCE, 246  
VERN L. BULLOUGH, 1636  
SPENCER E. CAHILL, 1356  
ROBERT B. CAIRNS, 1657  
RODERIC AI CAMP, 829  
JOHN L. CAMPBELL, 833  
JOSÉ CASANOVA, 247  
ALAN CAWSON, 256  
PATRICIA M. Y. CHANG, 553  
HARVEY M. CHOLDIN, 1091  
JAMES J. CHRISS, 1630  
PATRICIA TICINETO CLOUGH, 1073
- SAMUEL COHN, 275  
JAMES WILLIAM COLEMAN, 1094  
PETER W. COOKSON, 835  
L. L. CORNELL, 1340  
DIANA CRANE, 1362  
DANIEL M. CRESS, 1093  
COLIN CROUCH, 1349  
JANET CURRIE, 1660  
KATHLEEN DALY, 307  
ALISON DAVIS-BLAKE, 1662  
CARL N. DEGLER, 1079  
KEVIN J. DELANEY, 298  
TIA DENORA, 244  
NORMAN K. DENZIN, 558  
MARC EDELMAN, 1644  
STEPHEN ELLINGSON, 561  
BENJAMIN A. ELMAN, 817  
DORIS R. ENTWISLE, 839  
RICHARD V. ERICSON, 294  
THOMAS J. ESPENSHADE, 570  
GARY ALAN FINE, 1360  
SUE FISHER, 309  
RICHARD W. FLORY, 1064  
ANNE FONER, 1653  
MADELAINE GERBAULET-VANASSE, 271  
CALVIN GOLDSCHIEDER, 571  
RICHARD GRAHAM, 1637  
BRIAN GRAN, 1665  
WENDY GRISWOLD, 242, 1648  
PATRICIA A. GWARTNEY-GIBBS, 580
- PETER A. HALL, 1056  
DANIEL C. HALLIN, 549  
EUGENE HALTON, 1068  
ROGER HAMBURG, 263  
MASANORI HASHIMOTO, 253  
JEFFREY HAYDU, 1053  
HUGH HECLLO, 843  
PATRICK HELLER, 547  
GAIL HENDERSON, 301  
DONALD J. HERNANDEZ, 583  
ROSANNA HERTZ, 1336  
BARBARA HEYNS, 1343  
ALEXANDER HICKS, 543  
DAVID M. HUMMON, 1354  
HIROSHI ISHIDA, 1347  
ROBERT MAX JACKSON, 281  
JAMES M. JASPER, 260  
HARTMUT KAELEBLE, 252  
SATOSHI KANAZAWA, 1367  
DOUGLAS KELLNER, 1369  
ELAINE KEMPSON, 564  
DONGNO KIM, 267  
FRED KNISS, 1066  
PETER KOLLOCK, 831  
HANS PETER KRIESI, 249  
MICHÈLE LAMONT, 1069  
CHARLES C. LEMERT, 556  
TAK-CHUEN LUK, 828  
ANDREW MARCHANT-SHAPIRO, 265  
PHILIP A. MAY, 305  
SUSAN E. MAYER, 1655  
ALLAN MAZUR, 300  
DAVID S. MEYER, 261  
JOYA MISRA, 541  
JOHN MOHR, 277  
ALFONSO MORALES, 303  
CHANDRA MUKERJI, 814

## Book Reviewers

- |                              |                             |                                  |
|------------------------------|-----------------------------|----------------------------------|
| KEVIN NEUHOUSER,<br>1060     | HOWARD SCHUMAN, 289         | ABRAM DE SWAAN, 1055             |
| TANG NAH NG, 543             | MICHAEL SCHWARTZ,<br>1086   | RICHARD SWEDBERG,<br>563         |
| YUKO OGASAWARA, 1071         | DAVID O. SEARS, 1351        | ILAN TALMUD, 258                 |
| JANE M. PEDERSON, 279        | GAY W. SEIDMAN, 537         | JAY D. TEACHMAN, 585             |
| RICHARD A. PETERSON,<br>296  | ADAM B. SELIGMAN,<br>1365   | EDWARD A.<br>TIRYAKIAN, 1375     |
| FELIPE PIMENTEL, 269         | GI-WOOK SHIN, 826           | KEITH TRIBE, 821                 |
| STEPHEN B. PLANK, 837        | BARRY SMART, 1627           | MAMORU TSUKADA, 578              |
| KENNETH POMERANZ,<br>1646    | DAVID A. SMILDE, 1062       | STEPHEN TURNER, 1371             |
| RICHARD RUBINSON,<br>1058    | VICKI SMITH, 1346           | STEVEN PETER<br>VALLAS, 1084     |
| STEVEN RUGGLES, 532          | THOMAS SPENCE<br>SMITH, 554 | SUDHIR ALLADI<br>VENKATESH, 1632 |
| PAULA C. RUST, 851           | NADER SOHRABI, 1640         | DAVID G. WAGNER, 824             |
| PAUL A. SABATIER, 287        | AAGE B. SØRENSEN, 576       | MICHAEL WALLACE,<br>535          |
| COLIN SAMSON, 1080           | CHRIS SPARKS, 819           | JOHN R. WEEKS, 841               |
| CARL-ULRIK SCHIERUP,<br>1642 | J. WILLIAM SPENCER,<br>847  | MARTIN KING WHYTE,<br>568        |
| KATHRYN S. SCHILLER,<br>573  | VALERIE STEELE, 551         | JIM WOLF, 1364                   |
| MARK A. SCHNEIDER,<br>1629   | MARC W. STEINBERG,<br>848   | ALFORD A. YOUNG, JR.,<br>1089    |
|                              | ROBIN STRYKER, 292          |                                  |
|                              | JOHN R. SUTTON, 1096        |                                  |





# Columbia University Press

*New in paper...*

## Living on the Edge

The Realities of Welfare in America  
Mark Robert Rank, Washington University

260 PAGES / \$14.00, PAPER

"Rank skillfully combines ethnographic and statistical methods to portray the stark realities of life among welfare recipients, whose daily lives are driven by a struggle to survive. Event history methods are used to marshal 'hard' evidence that dispels myths about the reproductive and work behavior of welfare recipients, and challenges popular beliefs about the chronicity of welfare dependence. In-depth interviews with welfare recipients who represent a broad cross-section of the public assistance caseload bring to life the hardship, despair, and personal humiliation experienced by families forced to live 'on the edge.' His masterful combination of qualitative and quantitative evidence tells a more forceful story than either could alone."

—Marta Tienda, The University of Chicago



*With a new introduction by the author...*

## People, Plans, and Policies

Essays on Poverty, Racism, and Other National Urban Problems

Herbert J. Gans, Columbia University

412 PAGES / \$19.50, PAPER

Gans demonstrates that the basic urban problems are poverty and racism, and that until these concerns are addressed by increasing racial equality, creating jobs, and bringing about other reforms, the generally low quality of urban life will persist

*New...*

## The Dream of a New Social Order

Popular Magazines in America, 1893-1914

Matthew Schneirov, Duquesne University

368 PAGES / 20 ILLUS. / \$29.50, CLOTH

"Matthew Schneirov has given us a major breakthrough in social-history scholarship, because he takes seriously the trends of thought, or social consciousness, of 'mainstream' America and shows its powerful utopian, radical, and transformative characteristics. To miss this book is to miss a good chunk of twentieth-century America."

—Martin J. Sklar, Bucknell University

*New...*

## Circle of Protest

Political Ritual in the Tibetan Uprising, 1987-92

Ronald David Schwartz, Memorial University of Newfoundland

288 PAGES / \$16.50, PAPER / \$49.50, CLOTH

"Ronald David Schwartz has done a masterful job. Everyone interested in Tibet will want to read this fascinating book, both academics and others. It has real drama."

—James D. Seymour, Columbia University

*New in paper...*

## Class, Politics, and Ideology in the Iranian Revolution

Mansoor Moaddel, Eastern Michigan University

346 PAGES / \$17.50, PAPER

"A distinctive and valuable contribution ...

A comprehensive interpretation of the antecedents, underlying conditions, and development of the revolution that offers some sociological insights that will be fresh to readers."

—Richard Bullett, Columbia University

## Green Networks

A Structural Analysis of the Italian Environmental Movement

Mario Diani, University of Pavia and Bocconi University, Milan

Edinburgh University Press

Environment, Politics, and Society Series

300 PAGES / \$75.00, CLOTH

Using a systematic application of network analysis, Diani illuminates the social basis of collective political action, giving a fresh perspective on the meteoric rise of the environmental movement in Italian politics.

AT BETTER BOOKSTORES.



CREDIT CARDS ACCEPTED  
DEPT. S72  
136 SOUTH BROADWAY  
IRVINGTON, NY 10533  
TEL. (800) 944-8648  
FAX (800) 944-1844

CALL OR WRITE FOR OUR CATALOG.

# SOCIOLOGY CALIFORNIA

R W CONNELL

## **Masculinities**

Connell provides a nuanced and incisive analysis of how our notions of masculinity have evolved in psychoanalysis, social science, and historically in the creation of a global economy.  
\$40.00 cloth, \$16.00 paper

CAROL BROOKS GARDNER

## **Passing By**

*Gender and Public Harassment*  
"A unique study that will be a paradigm for others. . . Its contributions to the sociology of everyday life and to the understanding of public encounters and harassment are unparalleled."

—Douglas Maynard, University of Indiana, Bloomington  
\$40.00 cloth, \$16.00 paper

CHRISTENA L. TURNER

## **Japanese Workers in Protest**

*An Ethnography of Consciousness and Experience*

"An elegant contribution to our understanding of Japanese society. . . A model of clarity."

—Robert J. Smith, Cornell University  
\$35.00 cloth

At bookstores or order 1-800-822-6657.

## **NEW IN PAPERBACK**

JOHN A. CLAUSEN

## **American Lives**

*Looking Back at the Children of the Great Depression*  
"[A] heartwarming, richly detailed journey . . . An uplifting tribute to the sturdiness and strength of the human character."  
—*San Francisco Chronicle Review*

\$18.00 paper

ROB KLING, SPENCER OLIN, and MARK POSTER, Editors

## **Postsuburban California**

*The Transformation of Orange County since World War II*  
"A winning collection of essays about a county many believe to be paradigmatic of America's suburban future."

—*Los Angeles Times*  
\$15.00 paper

CLARENCE Y. H. LO

## **Small Property versus**

## **Big Government**

*Expanded & Updated Edition*  
"A compelling account of the California tax revolt and movements like it. . . A very good book."

—*Contemporary Sociology*  
\$14.00 paper

UNIVERSITY OF CALIFORNIA PRESS

**Education and Society** is an established international journal which publishes original research and contributions dealing with major up-to-date issues concerning the relationship between education and society. It provides a new forum, where teachers and scholars throughout the world, working in education and related areas, are able to examine and evaluate current issues and problems in education and society - from a broad and balanced social, cultural and economic perspective.

The journal should be of interest to teachers, academics, research scholars, educational administrators, graduate students and social workers.

Contributions are invited (normally no more than 5,000 words, typed double spaced in two copies) and should be addressed to the Editor, *Education and Society*, James Nicholas Publishers, P.O. Box 244, Albert Park, Australia, 3206.

In order to ensure that articles are of highest quality, all contributions are submitted to at least two independent referees before acceptance for publication.

The journal appears twice a year, in June and December. These two issues constitute one volume. An annual index is included in the December issue.

1995, Volume 13, ISSN 0726-2655

Subscription orders and requests for details should be addressed to:

JAMES NICHOLAS PUBLISHERS,  
P.O. Box 244, ALBERT PARK,  
VICTORIA, AUSTRALIA, 3206.

#### EDITOR

Dr Joseph Zajda, Australian Catholic University

#### CONSULTING EDITORS

Prof. Orlando Albornoz, University of Venezuela

Dr. Lilia Babaeva, Russian Academy of Sciences, Moscow

Prof. John Eggleston, University of Warwick

Prof. Nigel Grant, University of Glasgow

Prof. Henry Levin, Stanford University

Prof. Zola Malkova, Russian Academy of Education

Professor David Nyberg, State University of New York (Buffalo)

Prof. Douglas Sloan, Columbia University

Prof. Seth Spaulding, University of Pittsburgh

Dr. Joan Whitehead, University of Cambridge

Prof. Charles Willie, Harvard University

Dr. John Wilson, University of Oxford

#### RECENT AND FORTHCOMING ISSUES

Edward Berman THE POLITICAL ECONOMY OF AMERICAN EDUCATION REFORM IN THE 1990'S

Johann LeRoux & M E Smit THE ANTI-CHILD NATURE OF MODERN SOCIETY

Henry Giroux BORDER PEDAGOGY AND THE POLITICS OF POSTMODERNISM

George Psacharopoulos ON QUALITY OF EDUCATION

Kazim Bacchus IMPROVING THE QUALITY OF BASIC EDUCATION

John Martin Rich ASSESSMENT AND CONTROL OF SCHOOL VANDALISM

Martin Carnoy EDUCATION AND THE THEORIES OF THE STATE

Edmund King CHANGING EDUCATION FOR A CHANGING SOCIETY

Rea Zajda & Joseph Zajda MARKETING SCHOOL QUALITY AND EFFECTIVENESS IN A GLOBAL CULTURE

# EDUCATION AND SOCIETY

#### ORDER FORM

☐ Please enter our subscription to *Education and Society*

☐ We enclose \$ .....

☐ Please charge ☐ MasterCard ☐ Visa ☐ American Express Card No. ....

Expiry Date.....

Name.....

Address .....

Signed ..... Date .....

---

# **SOCIAL POLICY IN THE UNITED STATES**

Future Possibilities in  
Historical Perspective

**Theda Skocpol**

*"Social Policy in the United States bears out Theda Skocpol's reputation as the pre-eminent scholar of the history of social policymaking in the United States."*

—William Julius Wilson

*"Theda Skocpol has become the sociological analyst of the sources of American social policy. Her new book is must reading for anyone who seeks to understand the way in which we are different from Europe."*

—Seymour Martin Lipset

In this wide-ranging collection of essays, Theda Skocpol shows how historical understanding, centered on U.S. governmental institutions and shifting political alliances, can illuminate the limits and possibilities of American social policymaking both past and present.

Readers will be surprised at many of the findings of this volume. For example, Skocpol dispels the myth that Americans are inherently hostile to governmental social spending. When universal social programs jointly benefit the middle class and the poor, she shows, Americans since the nineteenth century have been willing to pay taxes for them and happy to partake of the security they provide. Skocpol's insights from the past illuminate many other issues of our time.

Cloth: \$29.95 ISBN 0-691-03786-8



## **ONE FOR ALL**

The Logic of Group Conflict

**Russell Hardin**

*"Russell Hardin develops a nice counterblast to those who think ethnic conflicts are somehow primordial and irrational."*

—John Dryzek

In a book that challenges the most widely held ideas of why individuals engage in collective conflict, Russell Hardin offers a timely, crucial explanation of group action in its most destructive forms. Contrary to those observers who attribute group violence to irrationality, primordial instinct, or complex psychology, Hardin uncovers a systematic exploitation of self-interest in the underpinnings of group identification and collective violence.

Exploring conflict between groups such as Serbs and Croats, Hutu and Tutsi, Northern Irish Catholics and Protestants, Hardin vividly illustrates the danger that arises when individual and group interests merge.

Cloth: \$24.95 ISBN 0-691-04350-7

---

## **PRINCETON UNIVERSITY PRESS**

AVAILABLE AT FINE BOOKSTORES OR DIRECTLY FROM THE PUBLISHER: 800-777-4726

---

## New Communitarian Thinking

*Persons, Virtues, Institutions,  
and Communities*  
Edited by AMITAI ETZIONI

Communitarian thought is at the heart of a fierce debate in political theory about the justice, efficacy, and future of liberalism and liberal societies. Amitai Etzioni has collected a sterling list of contributors who bring communitarian thinking to bear on such timely and contentious issues as abortion, homosexuality, free speech, and personal autonomy.

### Contributors

- Jeffrey Abramson, *Brandeis University*
- Elizabeth Bussette, *University of Massachusetts, Boston*
- Jean Bethke Elshtain, *University of Chicago*
- Robert Booth Fowler, *University of Wisconsin-Madison*
- Amy Gutmann, *Princeton University*
- David Hollenbach, *S.J., Boston College*
- Michael J. Sandel, *Harvard University*
- Philip Selznick, *University of California, Berkeley*
- Rogers M. Smith, *Yale University*
- Thomas A. Spragens, Jr., *Duke University*
- William M. Sullivan, *La Salle University*
- Charles Taylor, *McGill University*
- Daniel J. Tichenor, *Gordon Center at Brandeis University*
- Michael Walzer, *Institute for Advanced Study, Princeton*
- Alan Wolfe, *Boston University*

\$55.00 cloth, \$18.50 paper

UNIVERSITY PRESS  
OF VIRGINIA

Box 5008 University Station  
Charlottesville, VA 22903-0008



THIS  
PUBLICATION  
AVAILABLE  
FROM UMI

This publication is available from UMI in one or more of the following formats

- In Microform—from our collection of over 18,000 periodicals and 7,000 newspapers
- In Paper—by the article or full issues through UMI Article Clearinghouse
- Electronically, on CD-ROM, online, and/or magnetic tape—a broad range of ProQuest databases available including

abstract-and-index, ASCII full-text, and innovative full-image format

Call toll-free 800-521-0600, ext 2888, for more information, or fill out the coupon below

Name \_\_\_\_\_

Title \_\_\_\_\_

Company/Institution \_\_\_\_\_

Address \_\_\_\_\_

City/State/Zip \_\_\_\_\_

Phone ( ) \_\_\_\_\_

I'm interested in the following title(s) \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

UMI  
A Bell & Howell Company  
Box 49  
300 North Zeeb Road  
PO Box 1346  
Ann Arbor, MI 48106-1346  
800-521-0600 toll-free  
313-761-1203 fax

UMI



## THE ELOQUENCE OF SILENCE

Algerian Women in Question

Marnia Lazreg

"This book articulates in an authentic voice some of the major concerns of Third World women who, while committed to feminism, are deeply troubled by the Western, middle-class face of the movement. It is a book everyone should read."

—Azizah al-Hibri, University of Richmond

288 pp \$16.95/Pb

Second Edition

## ETHNOGRAPHY

Principles in Practice

Paul Atkinson

and Martyn Hammersley

Thoroughly updated and substantially rewritten, the second edition of this popular textbook is now even more relevant and useful for students and researchers. In this accessible introduction to the methods of ethnographic fieldwork, Hammersley and Atkinson reconsider the status of ethnography and seek to place it quite explicitly in a general methodological context.

336 pp \$18.95/Pb

## CONCEIVING SEXUALITY

Approaches to Sex Research in a Postmodern World

Edited by Richard G. Parker and John H. Gagnon

Edited by two of the leading figures in contemporary sex research, *Conceiving Sexuality* brings together the contributions of writers from a wide range of social science disciplines and cultural traditions who are working at the cutting edge of contemporary sex research.

320 pp \$18.95/Pb

## SOCIAL EXPERIENCE AND ANTHROPOLOGICAL KNOWLEDGE

Edited by Kirsten Hastrup and Peter Hervik

*Social Experience and Anthropological Knowledge* traces the process from the ethnographic experience to the analytical results, showing how fieldwork enables the ethnographer to arrive at an understanding, not only of "culture" and "society," but also of the processes by which cultures and societies are transformed.

224 pp \$17.95/Pb

## WHITE RACISM

The Basics

Joe R. Feagin and Hernán Vera

The authors analyze a series of racial incidents—public and private—which reveal the "dirty little secret of American life" of white-on-black racism as a social problem deeply embedded in culture, the law, crime and punishment, the shape of cities and the environment, international relations, and daily rhythms of life.

224 pp \$16.95/Pb

# Theorizing Society

ANNOUNCING A NEW SERIES

## MAIN TRENDS OF THE MODERN WORLD

EDITED BY ROBERT JACOBSON AND ALAN SICA

*"Should become a staple for scholars in all the social sciences...a most worthy publishing undertaking and event."*

—Alan Sica, Editor, *Sociological Theory*

This bold and innovative series addresses the defining characteristics of our world in a cohesive and constructive fashion, while analyzing the main trends and experiences of the modern world. Each volume in the series focuses on one decisive aspect of modern life and features the best classical and contemporary essays on that subject, representing social and political theory at its finest.

### PROPAGANDA

EDITED BY ROBERT JACOBSON

Traces the origins, ethos, and workings of modern propaganda.

480 pages

ISBN 0-8146-4197-1 \$37.50 paper

### THE NEW MIDDLE CLASSES

LIFE STYLES, STATUS CLAIMS, AND POLITICAL ORIENTATIONS

EDITED BY ALAN SICA

Explores the origins of the new American middle classes that emerged in the 20th century and reveals the relationship of these classes to capitalism, bureaucracy, and politics.

480 pages

ISBN 0-8146-4220-1 \$37.50 paper

### EAST GERMAN DISSIDENTS AND THE REVOLUTION OF 1989

SOCIAL MOVEMENT IN A LENINIST REGIME

EDITED BY ALAN SICA

*"Deserves to be a 'classic' in political sociology...All the ingredients for a truly exceptional book."*

—Andreas S. Markovits

University of California, Santa Cruz

360 pages

ISBN 0-8146-4219-4 \$37.50 paper

### METROPOLIS

CENTER AND SYMBOL OF OUR TIMES

EDITED BY FRITZ KASSEL

Drawing on social thinkers such as Georg Simmel, Lewis Mumford, and Herbert Marcuse, *Metropolis* chronicles the rise of the modern city.

480 pages

ISBN 0-8146-4198-9 \$37.50 paper

### SOCIAL MOVEMENTS

CRITIQUES, CONCEPTS, CASE STUDIES

EDITED BY STANISLAV M. LEVAK

Provides a well-rounded and provocative perspective on this most compelling form of political expression.

480 pages

ISBN 0-8146-4218-6 \$37.50 paper

### POVERTY AND THE UNDERCLASS

CHANGING PERCEPTIONS OF THE POOR IN AMERICA

WILLIAM A. FELSO

Offering new explanation for the persistence of poverty, William A. Felsos looks beyond familiar labels such as "underclass" and "family values" to analyze poverty not as a monolithic condition, but as a multi-dimensional problem.

360 pages

ISBN 0-8146-4218-6 \$37.50 paper



## NEW YORK UNIVERSITY PRESS

70 Washington Square South • New York, NY 10012

1-800-996-NYUP (6987) • Fax: (212) 996-3833

800



0002-9602(199505)100:6;1-Z